

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 01:53:16 ; Search time 135 Seconds  
(without alignments)  
732.298 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHLLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1214	100.0	225	3 AAY99836	Aay99836 Discosoma
2	1214	100.0	225	3 AAB01622	Aab01622 Discosoma
3	1214	100.0	225	4 AAG65509	Aag65509 Anthozoa
4	1214	100.0	225	5 ABB08834	Abb08834 Yeast opt
5	1214	100.0	225	5 AAE28833	Aae28833 Discosoma
6	1214	100.0	225	5 AAE17540	Aae17540 Discosoma
7	1214	100.0	225	5 AAO18270	Aao18270 Discosoma
8	1214	100.0	225	5 ADY51716	Ady51716 Discosoma
9	1214	100.0	225	6 AAE34962	Aae34962 Discosoma
10	1214	100.0	225	7 ADC24126	Adc24126 Discosoma
11	1214	100.0	225	7 ABW00918	Abw00918 Discosoma
12	1214	100.0	225	7 ADP70403	Adp70403 Discosoma
13	1214	100.0	225	7 ADH34489	Adh34489 Discosoma
14	1214	100.0	225	7 ADL46203	Adl46203 Discosoma
15	1214	100.0	225	7 ADN33979	Adn33979 Wild-type
16	1214	100.0	225	8 ADI36421	Adi36421 Discosoma
17	1214	100.0	225	8 ADM97769	Adm97769 D sp red
18	1214	100.0	225	8 ADQ28780	Adq28780 Jellyfish
19	1214	100.0	225	9 ADX26534	Adx26534 Discosoma
20	1214	100.0	487	5 ABB08821	Abb08821 Autofluor
21	1214	100.0	506	5 ABB08822	Abb08822 Autofluor
22	1214	100.0	547	5 ABB08823	Abb08823 Autofluor
23	1211	99.8	225	5 AAE28920	Aae28920 Discosoma
24	1211	99.8	225	7 ABW00937	Abw00937 Discosoma

25	1211	99.8	225	7 ABW00929	Abw00929 Discosoma
26	1211	99.8	225	7 ABW00938	Abw00938 Discosoma
27	1211	99.8	225	7 ADH34498	Adh34498 Discosoma
28	1211	99.8	225	7 ADH34499	Adh34499 Discosoma
29	1210	99.7	225	5 AAE28919	Aae28919 Discosoma
30	1210	99.7	225	7 ABW00930	Abw00930 Discosoma
31	1210	99.7	226	4 AAG65510	Aag65510 Anthozoa
32	1210	99.7	242	7 ADE24109	Ade24109 Discosoma
33	1210	99.7	545	7 ADL18132	Adl18132 RFP:PS (NI
34	1210	99.7	548	7 ADL18156	Adl18156 RFP:PS (HI
35	1209	99.6	225	5 AAE28922	Aae28922 Discosoma
36	1209	99.6	225	7 ABW00936	Abw00936 Discosoma
37	1209	99.6	225	7 ABW00931	Abw00931 Discosoma
38	1209	99.6	226	5 ABB08835	Abb08835 Yeast opt
39	1208	99.5	225	7 ABW00932	Abw00932 Discosoma
40	1208	99.5	225	7 ABW00935	Abw00935 Discosoma
41	1208	99.5	225	7 ABW00939	Abw00939 Discosoma
42	1208	99.5	225	7 ADH34500	Adh34500 Discosoma
43	1207	99.4	225	5 AAE28921	Aae28921 Discosoma
44	1207	99.4	225	5 AAE17541	Aae17541 Discosoma
45	1207	99.4	225	5 ADY51734	Ady51734 Discosoma

ALIGNMENTS

RESULT 1  
AAY99836  
ID AAY99836 standard; protein; 225 AA.

AC AAY99836;  
XX  
XX  
DT 12-SEP-2003 (revised)  
DT 19-SEP-2000 (first entry)  
XX  
DE Discosoma sp. "red" novel fluorescent protein drFP583.  
XX  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
XX  
XX  
OS Discosoma sp; "red".  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 122 /note= "encoded by TC"  
FT Misc-difference 127 /note= "encoded by GTTG"  
XX  
XX  
PN WO200034326-A1.

XX  
PD 15-JUN-2000.  
XX  
XX  
PF 10-DEC-1999; 99WO-US029473.  
XX  
PR 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
DR WPI; 2000-423381/36.  
XX  
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
PS Claim 20; Page 74-75; 86pp; English.  
XX  
CC The present sequence is a novel fluorescent protein (nFP) encoded by the  
CC full-length cDNA drFP583. drFP583 was isolated from Discosoma sp. "red",  
CC a non-bioluminescent species of the Class Anthozoa. Fluorescent proteins  
CC can be used in fluorescent labeling, a useful tool for marking a protein,







CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents an anthozoan fluorescent protein

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGSPLPFAMDI 60  
|||  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGSPLPFAMDI 60  
DB 1 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
|||  
61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
|||  
121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSI 180  
|||  
121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSI 180  
OY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB

RESULT 4

ABB08834  
ID ABB08834 standard; protein; 225 AA.

XX AC ABB08834;

XX DT 29-MAY-2002 (first entry)

XX DE Yeast optimised RFP SEQ ID NO 17.

XX KM Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
XX KM Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
XX KM Escherichia coli; green fluorescent protein; biotechnology.

XX OS Anthozoa.

XX PN DE20001395-U1.

XX PD 15-MAR-2001.

XX PF 27-JAN-2000; 2000DE-02001395.

XX PR 27-JAN-2000; 2000DE-02001395.

XX PA (GPCB-) GPC BIOTECH AG.

XX DR WPI; 2002-228394/29.

XX DR N-PSDB; ABA95905, ABA95921, ABA95922.

XX PT New DNA encoding red fluorescent protein, useful as marker in  
XX PT biotechnology, has sequence optimised for expression in eukaryotes,  
XX PT especially yeast or plants.

XX PS Disclosure; Page 13-14; 19pp; German.

XX CC The invention relates to DNA (I) containing either sequence ABA95905 or  
XX CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein

CC (YFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
XX sequence is that of the yeast optimised RFP

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGSPLPFAMDI 60  
|||  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGSPLPFAMDI 60  
DB 1 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
|||  
61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERVMNFBEDGCVTVTQDSSLQDGCFTY 120  
|||  
121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSI 180  
|||  
121 KYKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEFKSI 180  
OY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 5

AAE28833  
ID AAE28833 standard; protein; 225 AA.

XX AC AAE28833;

XX DT 27-DEC-2002 (first entry)

XX DE Discosoma sp. drFP583 (NFP-6) wild-type protein.

XX KM Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
XX KM fluorescence activated cell sorting application; fluorescent timer;  
XX KM biosensor; fluorescence resonance energy transfer application; FRET;  
XX KM colouring agent; recombinant DNA application; analyte detection assay;  
XX KM sunscreen; second messenger detector; drFP583 protein; NFP-6;

XX OS Discosoma sp.

XX PN WO200268459-A2.

XX PD 06-SEP-2002.

XX PF 20-FEB-2002; 2002WO-US005749.

XX PR 21-FEB-2001; 2001US-0270983P.

XX PR 04-DEC-2001; 2001US-00006922.

XX PA (CLON-) CLONTECH LAB INC.

XX PI Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;

XX DR WPI; 2002-691654/74.

XX DR N-PSDB; AAD46278.

XX PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
XX PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
XX PT analyte detection assays or fluorescence activated cell sorting  
XX PT applications.



PS Disclosure; Page 70-71; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drfp583 (NRF-6) wild-type protein of the  
CC invention

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180

QY 181 YMAKKPVQLPGYYVVD SKLDTITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVD SKLDTITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 6  
AAE17540  
ID AAE17540 standard; protein; 225 AA.  
XX  
AC AAE17540;  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. humanised wild-type Anthozoa protein drfp583.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; anthozoa protein; drfp583.  
XX  
OS Discosoma sp.  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Pradkov AF, Terakikh A;  
XX  
OS  
XX  
DR WPI; 2002-154595/20.  
DR N-PSDB; AAD28207.  
XX  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
PS Example 1; Fig 1; 89pp; English.

XX The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodelling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated proteins, and  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigations where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is Discosoma sp.  
CC humanised wild-type Anthozoa protein drfp583 used for generating  
CC fluorescent proteins

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180

QY 181 YMAKKPVQLPGYYVVD SKLDTITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVD SKLDTITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 7  
AAO18270  
ID AAO18270 standard; protein; 225 AA.  
XX  
AC AAO18270;  
DT 26-SEP-2002 (first entry)  
XX  
DE Discosoma red fluorescent protein.  
XX  
KW Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
KW modified yeast strain; environmental pollution.  
XX  
OS Discosoma sp.  
XX  
PN DE10061872-A1.  
XX  
PD 20-JUN-2002.  
XX  
PF 12-DEC-2000; 2000DE-01061872.  
XX  
PR 12-DEC-2000; 2000DE-01061872.  
XX



PA (LICH/) LICHTENBERG-FRATE H.  
XX Lichtenberg-Frate H;  
PI  
XX WPI; 2002-539633/58.  
DR N-PSDB; AAL47952.  
XX  
PT Modified yeast strain, useful for detecting toxic compounds in  
PT environment, contains integrated cassettes responsive to genotoxic and  
PT cytotoxic compounds.  
XX  
PS Disclosure; Page 21-22; 34pp; German.  
XX  
XX The present invention relates to a modified yeast strain that contains,  
CC integrated stably and functionally in its genome, a genotoxicity cassette  
CC and a cytotoxicity cassette, each comprising a promoter and reporter  
CC gene, both of which are different in the two cassettes. The modified  
CC yeast strain is used to detect environmental pollution, especially  
CC genotoxic and/or cytotoxic substances in complex environmental  
CC contaminants, especially organic compounds, but also (non-)ionising  
CC radiation and chemical carcinogens. Particular applications are in  
CC monitoring (waste) water (e.g. as an early warning system), medical  
CC toxicology screening and for industrial process control. The present  
CC sequence is a marker protein suitable for use in the cassettes of the  
CC present invention  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 60  
QY 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGCVTVTQDSSLQDGCFTY 120  
QY 121 KVKEIGVNFPSDGPVMOKKTMGWEASTERLPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
DB 121 KVKEIGVNFPSDGPVMOKKTMGWEASTERLPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 8  
ADY51716  
ID ADY51716 standard; protein; 225 AA.  
XX  
AC ADY51716;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE Discosoma sp. wild type red fluorescent protein (RFP) Seq 12.  
XX  
KW fluorescence; mutagenesis; red fluorescent protein; protein interaction.  
XX  
OS Discosoma sp.  
XX  
PN WO200268605-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 26-FEB-2002; 2002WO-US006063.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (REGC ) UNIV CALIFORNIA.

XX Tsien RY, Baird GS, Campbell RE, Zacharias DA;  
PI  
XX WPI; 2002-713372/77.  
DR N-PSDB; ADY51715.  
XX  
PT New non-oligomerizing fluorescent protein containing at least one  
PT mutation that reduces or eliminates the ability of the protein to  
PT oligomerize, useful for making better and new assays for molecular  
PT biology.  
XX  
PS Claim 10; SEQ ID NO 12; 117pp; English.  
XX

CC This invention relates to a novel non-oligomerizing fluorescent protein.  
CC Specifically, it refers to the presence of at least one mutation in the  
CC fluorescent protein that reduces or eliminates the ability of the protein  
CC to oligomerize. The present invention describes fluorescent proteins and  
CC derived from naturally occurring green or red fluorescent proteins and  
CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
CC protein linked to at least one protein of interest. As such, these fusion  
CC proteins can be used in methods and compositions to determine the pH of a  
CC sample, or whether the sample contains an enzyme, molecule or agent that  
CC regulates the activity of an expression control sequence. Furthermore,  
CC they may be used to identify a specific interaction of molecules, such  
CC that they are useful for improving or developing new assays in the field  
CC of molecular biology. This polypeptide sequence is the Discosoma sp. wild  
CC type red fluorescent protein (RFP) of the invention.  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 60  
QY 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGCVTVTQDSSLQDGCFTY 120  
QY 121 KVKEIGVNFPSDGPVMOKKTMGWEASTERLPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
DB 121 KVKEIGVNFPSDGPVMOKKTMGWEASTERLPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 9  
AAE34962  
ID AAE34962 standard; protein; 225 AA.  
XX  
AC AAE34962;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Discosoma species red fluorescent protein (RFP).  
XX  
KW phosphorylation indicator; fluorescent protein; detection; phosphatase;  
XX  
OS Discosoma sp.  
XX  
PN WO200295058-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 24-MAY-2002; 2002WO-US016955.  
XX  
PR 24-MAY-2001; 2001US-00865291.  
XX



XX (REGC ) UNIV CALIFORNIA.  
PA  
XX  
PI Tsien RY, Ting AY, Zhang J;  
XX  
DR WPI: 2003-148474/14.  
XX N-PSDB; AAD53432.  
PT  
PT Novel chimeric phosphorylation indicators, useful for detecting  
PT kinase/phosphatase in samples, has donor molecule, phosphorylatable  
PT domain, phosphoaminoacid binding domain, and acceptor molecule, in  
PT operative linkage.  
XX  
XX  
PS Disclosure; Col 65-66; 38pp; English.  
XX  
CC The present invention relates to chimeric phosphorylation indicators  
CC comprising a phosphorylation polypeptide and a fluorescent protein or in  
CC operative linkage, a donor molecule, a phosphorylatable domain, a  
CC phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The  
CC phosphorylation indicators of the invention are useful for detecting  
CC kinases or phosphatases in a biological sample. They are also useful in  
CC high throughput analysis e.g. for detecting a kinase inhibitor or  
CC phosphatase inhibitor. The present sequence is Discosoma species red  
XX fluorescent protein (RFP) used in the invention  
SQ Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 6;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MSSKQVNIKEFMRFKVRMEGTANGHEFEIIEGEGEGRPYEGHNTVKLKVTIKGGGLPFAWDI	60
Db	1	MSSSKVNIKEFMRFKVRMEGTANGHEFEIIEGEGEGRPYEGHNTVKLKVTIKGGGLPFAWDI	60
Qy	61	LSPQFOYGSKVYVKHPADIPDYKKLSPEEGFKMERVMNFEDGCVTVTQDSIQDGCFIY	120
Db	61	LSPQFOYGSKVYVKHPADIPDYKKLSPEEGFKMERVMNFEDGCVTVTQDSIQDGCFIY	120
Qy	121	KVKEIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI	180
Db	121	KVKEIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI	180
Qy	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL	225
Db	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL	225

```

RESULT 10
ADC24126
ID ADC24126 standard; protein; 225 AA.
XX
AC ADC24126;
XX
DT 18-DEC-2003 (first entry)
XX
DE Discosoma wild-type red fluorescent protein.
XX
KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;
KW fluorescent protein variant; transcription induction detection;
KW fluorescence energy resonance transfer; FRET; protein kinase;
KW protein phosphatase; ion indicator.

```

XX (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
PA  
XX  
PI Tsien RY, Campbell RE;  
XX  
DR WPI; 2003-743764/70.  
DR N-PSDB; ADC24127, ADC24134.  
XX  
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.

Query Match	100.0%;	Score 1214;	DB 7;	Length 225;
Best Local Similarity	100.0%;	Pred. NO. 1e-127;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 11
ABW000918
ID      ABW000918  standard; protein; 225 AA
XX
AC      ABW000918;

```



XX 15-JAN-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein (RFP).  
DE  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP.  
XX  
XX Discosoma sp.  
OS  
XX US2003170911-A1.  
XX  
XX 11-SEP-2003.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
XX  
XX (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
XX Telen RY, Zacharias DA, Baird GS;  
PI  
XX WPI; 2003-802418/75.  
DR N-PSDB; AAD61969.  
XX  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
XX Claim 10; Page 30-31; Opp; English.  
PS  
XX The invention relates to a non-oligomerizing fluorescent protein  
XX containing a mutation that reduces or eliminates its ability to  
CC oligomerize. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP)  
XX  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHHYLVFPKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHHYLVFPKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 12  
ADP70403  
ID ADP70403 standard; protein; 225 AA.  
XX  
XX ADF70403;  
AC  
XX  
DT 12-FEB-2004 (first entry)  
XX

DE Discosoma wild-type GFP variant protein SeqID26.  
XX  
XX ligand; orphan receptor protein; fusion protein; fluorescent protein;  
KM cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KM GFPuv; Enhanced GFP; EGFP.  
XX  
XX Discosoma sp.  
OS  
XX WO2003071272-A1.  
XX  
XX 28-AUG-2003.  
XX  
XX 21-FEB-2003; 2003WO-JP001901.  
XX  
XX 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosooya M;  
PI  
XX WPI; 2003-697654/66.  
DR N-PSDB; ADF70404.  
XX  
XX Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
XX Disclosure; SEQ ID NO 26; 594pp; Japanese.  
PS  
XX This invention relates to a novel method of identifying ligands to an  
XX orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein for incorporation in  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFBEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHHYLVFPKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHHYLVFPKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225

RESULT 13  
ADH34489  
ID ADH34489 standard; protein; 225 AA.  
XX  
XX ADH34489;  
AC  
XX  
DT 11-MAR-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein DsRed (wild-type).



XX Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labeling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
XX red fluorescent protein; DsRed.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT MISC-difference 146 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 148"  
FT MISC-difference 161 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 165"  
FT MISC-difference 163 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 167"  
FT MISC-difference 197 /note= "Key position for chromoprotein/fluorescent  
FT protein activity. Corresponds to GFP residue 203"  
XX WO2003057833-A2.  
XX  
PD 17-JUL-2003.  
XX  
PF 23-DEC-2002; 2002WO-US041418.  
XX  
PR 26-DEC-2001; 2001US-0343128P.  
XX (CLON-) CLONTECH LAB INC.  
PA  
XX  
PI Bulina ME, Chudakov D, Lukyanov KA;  
XX  
DR WPI; 2003-607998/57.  
XX  
PT Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
PS Example 1; Fig 1; 56pp; English.  
XX  
CC The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs,  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as colouring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and  
CC are also useful as sunscreens and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect

CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a wild-type red fluorescent protein, DsRed, from Discosoma sp.  
CC that was used as a parent sequence for the generation of mutant proteins  
CC in an example of the invention.  
XX  
SQ Sequence 225 AA;  
XX  
Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
QY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGCVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMOQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMOQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 14  
ADL46203  
ID ADL46203 standard; protein; 225 AA.  
XX  
AC ADL46203;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma red fluorescent (DsRed) protein.  
XX  
KW red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.  
XX  
PN WO2003086446-A1.  
XX  
PD 23-OCT-2003.  
XX  
PF 09-APR-2003; 2003WO-US010879.  
XX  
PR 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Campbell RE, Baird GS;  
XX  
DR WPI; 2003-845265/78.  
DR N-PSDB; ADL46204.  
XX  
PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
PS Claim 1; SEQ ID NO 1; 166pp; English.



XX The invention relates to a polynucleotide sequence encoding a *Discosoma*  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunosays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC protein.

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

Db 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY 61 LSPQFQYGSKVYVYKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120

Db 61 LSPQFQYGSKVYVYKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 15

ADN33979  
ID ADN33979 standard; protein; 225 AA.

XX ADN33979;

AC 01-JUL-2004 (first entry)

DT wild-type DsRed protein.

DE Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRed.

XX *Discosoma* sp.

OS WO2003054158-A2.

XX 03-JUL-2003.

XX 18-DEC-2002; 2002WO-US040539.

XX 19-DEC-2001; 2001US-0341723P.

XX (UYCH-) UNIV CHICAGO.

XX Bevis B, Glick B;

XX WPI, 2003-569236/53.

DR N-PSDB; ADN33978.

XX Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent

PT mutant of a Cnidarian chromo- or fluorescent protein or its mutant,

PT useful for applications involving chromo- or fluorescent proteins.

XX Claim 8; SEQ ID NO 2; 65pp; English.

CC The present invention relates to nucleic acid that encodes a rapidly  
CC maturing chromo or fluorescent mutant of a Cnidarian chromo- or  
CC fluorescent protein or its mutant. The protein is useful in applications  
CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
CC useful for producing a chromo and/or fluorescent protein which involves  
CC growing the cell, whereby the protein is expressed, and isolating the  
CC protein substantially free of other proteins. The protein is useful in  
CC applications involving chromo- or fluorescent protein and is useful as  
CC PCR primers, hybridization related probes, etc. The expression cassettes are  
CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for  
CC biological analytes of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins  
CC also find use in protease cleavage assays. The proteins can also be used  
CC in assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRed.

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

Db 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY 61 LSPQFQYGSKVYVYKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120

Db 61 LSPQFQYGSKVYVYKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

Search completed: January 11, 2006, 02:02:58  
Job time : 136 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 01:55:51 ; Search time 16 Seconds  
(without alignments)  
1353.048 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHLEFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211.5	17.4	238	1 JQ1514	green-fluorescent
2	93.5	7.7	26926	1 I38344	titin, cardiac mus
3	92	7.6	1873	2 T30944	surface protein pr
4	91	7.5	458	1 A42386	hsp 90-binding pro
5	89	7.3	1433	1 A36734	baclillopeptidase F
6	87.5	7.2	1116	2 B70476	hypothetical prote
7	87	7.2	568	2 T06489	probable peptidylp
8	87	7.2	1484	2 C97196	signal-transducin
9	85.5	7.0	456	2 G69397	hypothetical prote
10	85	7.0	271	2 F69442	adhesion-type prot
11	84.5	7.0	679	2 A40351	neural cell adhesi
12	84.5	7.0	725	1 TJMSNG	S-layer protein -
13	84	6.9	1616	2 T17884	conserved hypothet
14	83.5	6.9	268	2 E90276	hypothetical prote
15	83.5	6.9	340	2 B69544	gibberellin 3 beta
16	83.5	6.9	374	2 T06245	gibberellin 3 beta
17	83.5	6.9	374	2 T06244	gibberellin 3 beta
18	83	6.8	373	2 T50605	hypothetical prote
19	83	6.8	15281	2 S41309	cyclosporin synthe
20	82.5	6.8	559	1 S55383	peptidylprolyl iso
21	82.5	6.8	680	2 S17982	Kallmann syndrome
22	82.5	6.8	931	2 T32919	hypothetical prote
23	82	6.8	341	2 B53125	restriction enzyme
24	81.5	6.7	292	2 C69106	glucose-1-phosphat
25	81.5	6.7	296	2 I37989	La 4.1 protein - h
26	81.5	6.7	551	1 S72485	peptidylprolyl iso
27	81.5	6.7	629	2 C64180	hypothetical prote
28	81.5	6.7	6805	2 S20901	titin - rabbit (fr
29	81	6.7	346	2 S77025	nitrlase (EC 3.5.

30	81	6.7	587	1 E69171	phosphoesterase-re
31	80.5	6.6	404	1 S03849	ribonucleoprotein
32	80.5	6.6	415	1 JC1494	ribonucleoprotein
33	80.5	6.6	862	2 F75116	hypothetical prote
34	80.5	6.6	1003	1 AJCHPR	phosphoribosylamin
35	80.5	6.6	1115	1 TJMSNL	neural cell adhesi
36	80	6.6	621	2 A95250	choline binding pr
37	80	6.6	690	2 F98114	immunophilin FKBP5
38	79.5	6.5	459	2 A46372	conserved hypothet
39	79.5	6.5	475	2 T44566	hypothetical prote
40	79.5	6.5	487	2 T45982	ABC transporter (p
41	79.5	6.5	648	2 B84139	neural cell adhesi
42	79.5	6.5	853	1 IJBONC	immunophilin p59 -
43	79	6.5	458	2 JN0873	hypothetical prote
44	79	6.5	703	2 A64351	disease resistance
45	79	6.5	1214	2 T47438	

ALIGNMENTS

RESULT 1  
JQ1514  
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)  
C/Species: Aequorea victoria  
C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 31-Dec-2004  
C/Accession: J50692; JQ1514; FQ0335; S48693; S51330; S51331  
R/Prabher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.  
Gene 111, 229-233, 1992  
A/Title: Primary structure of the Aequorea victoria green-fluorescent protein.  
A/Reference number: JQ1514; MUID:92175527; PMID:1347277  
A/Accession: J50692  
A/Molecule type: DNA  
A/Residues: 1-107,'S',109-238 <PRA1>  
A/Cross-references: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; UNIPARC:UPI000016B7F4  
A/Accession: JQ1514  
A/Molecule type: mRNA  
A/Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>  
A/Cross-references: UNIPARC:UPI000002FBA; GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:5  
A/Accession: PQ0335  
A/Molecule type: protein  
A/Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>  
A/Cross-references: UNIPARC:UPI0000173FD6; UNIPARC:UPI0000173FD7; UNIPARC:UPI0000173FD8;  
R/Inouye, S.; Tsuji, F.I.  
FEBS Lett. 351, 211-214, 1994  
A/Title: Evidence for redox forms of the Aequorea green fluorescent protein.  
A/Reference number: S48693; MUID:94364470; PMID:8082767  
A/Accession: S48693  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>  
A/Cross-references: UNIPARC:UPI000003582F; GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:5  
R/Watkins, J.N.; Campbell, A.K.  
submitted to the EMBL Data Library, January 1995  
A/Reference number: S51330  
A/Accession: S51330  
A/Molecule type: mRNA  
A/Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',;  
A/Cross-references: UNIPARC:UPI000007BC01; EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PIT  
A/Experimental source: clone gfp1  
A/Accession: S51331  
A/Molecule type: mRNA  
A/Residues: 1-24,'Q',26-29,'R',31-83,'L',85-153,'G',155-156,'P',158-171,'K',173-208,'Q',;  
A/Cross-references: UNIPARC:UPI0000079E2D; EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PIT  
A/Experimental source: clone gfp2  
A/Reference number: A65692; PDB:1GFL  
A/Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A',2-79,'R',81-9;  
A/Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli  
R/Yang, F.; Moss, L.G.; Phillips Jr., G.N.  
Nat. Biotechnol. 14, 1246-1251, 1996  
A/Title: The molecular structure of green fluorescent protein.



A;Reference number: A58953; MUID:98294543; PMID:9631087  
A;Contents: annotation; X-ray crystallography, 1.9 angstroms  
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFN) emitting  
C;Genetics: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr  
A;Gene: GFP  
A;Introns: 69/3; 167/3  
C;Keywords: chromoprotein; luminescence  
F;65-67/Cross-link: 5-imidazolinone (Ser-Gly) #status experimental  
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match            17.4%; Score 211.5; DB 1; Length 238;  
Best Local Similarity   25.7%; Pred. No. 3.3e-11;  
Matches   56; Conservative   47; Mismatches   102; Indels   13; Gaps   6;

Db      16 VRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGEBLPFAWDILSPQFGYGSKYVKH 75  
       |:|:|||||:|:||||| |::| |::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db      16 VELDGDVNGHKFSVSGEGDATYGKLTlKFICTT--GKLVPWPPTLVTFSTYGVCFSRY 74  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Qy      76 PADIP--DYKKLSFPPEGFKMERVMNFEDGGVVTVTDSSSLQDGCFIYKKVFIGNPPSDG 133  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db      75 PDHMKQHDFEKSAMPEGYQERTIFYKDDGNKYTRAEVKEGDTLVNRIELKIGDKPEDG 134  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Qy      134 PWMOKTMGWESTERLYPRDGLKGEIHKALKL---DGHYLVEF--KSIYAKKPV 187  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db      135 NILGHK-MEYNYNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPV 193  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Qy      188 QLPGYYYVDSKLDIT---SHNEDYTIVEQYERTEGRHH 222  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db      194 LLPDNHYLSTQSALSKDPNEKRDMILLFVTAAGITH 231  
       |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:

A;Accession: S20899  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'Q', 22455-22480, 'TR', 22483-2;  
A;Cross-references: UNIPARC:UPI0000172571; EMBL:X64697; NID:g37190; PIDN:CAA45938.1; PID:  
R;Kolmerer, B.; Olivier, N.; Witt, C.C.; Herrmann, B.G.; Labelt, S.  
J. Mol. Biol. 256, 556-563, 1996  
A;Title: Genomic organization of M line titin and its tissue-specific expression in two c  
A;Reference number: S63665; MUID:96177761; PMID:8604138  
A;Accession: S63665  
A;Status: nucleic acid sequence not shown  
A;Molecule type: DNA  
A;Residues: 26729-26825 <KOL>  
A;Cross-references: UNIPARC:UPI0000172572; EMBL:X92412; NID:g1236761  
R;Gaute, M.; Leonard, K.; Labelt, S.  
EMBO J. 12, 3827-3834, 1993  
A;Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiat  
A;Reference number: S37393; MUID:94008990; PMID:8404852  
A;Accession: S37393  
A;Molecule type: mRNA  
A;Residues: 26831-26926 <GAU>  
A;Cross-references: UNIPARC:UPI0000172573  
R;Improta, S.; Politou, A.S.; Pastore, A.  
submitted to the Brookhaven Protein Data Bank, February 1996  
A;Reference number: A66736; PDB:1TIT  
A;Contents: annotation; conformation by (1)H-NMR, residues 5253-5341  
R;Pfuhl, M.; Pastore, A.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A;Reference number: A66201; PDB:1NCT  
A;Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155

RESULT 2

I38344  
titin, cardiac muscle [validated] - human

N:Alternate names: connectin  
N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.-)  
C:Species: Homo sapiens (man)  
C:Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: I38344; I38345; S20898; S20897; S20899; S63665; S37393  
R:Labelt, S.; Kolmerer, B.  
Science 270, 293-296, 1995

A:Title: Tittins: giant proteins in charge of muscle ultrastructure and elasticity.  
A:Reference number: A57430; MUID:96026330; PMID:7569978  
A:Accession: I38344

A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL

A:Molecule type: mRNA

A:Residues: 1-26926 <LAB1>  
A:Cross-references: UNIPROT:Q10466; UNIPARC:UPI000017CF2C; EMBL:X90568; NID:g1017424; PDB:1MUSCO, G.; Tziatzios, C.; Schuck, P.; Pastore, A.  
Biochemistry 34, 553-561, 1995

A:Title: Dissecting titin into its structural motifs: identification of an alpha-helix n  
A:Reference number: I38345; MUID:95119041; PMID:7819249  
A:Accession: I38345

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1977-2014 <MUS>  
A:Cross-references: UNIPARC:UPI000016B0E9; EMBL:X83270; NID:g602579; PIDN:CAA58243.1; PDB:1RIabelt, S.; Gautel, M.; Lakey, A.; Trinick, J.  
EMBO J. 11, 1711-1716, 1992

A:Title: Towards a molecular understanding of titin.  
A:Reference number: S20897; MUID:92258380; PMID:1582406  
A:Accession: S20898

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 13597-14200,'I',14202-14696 <LAB2>  
A:Cross-references: UNIPARC:UPI000016B0E7; EMBL:X64698; NID:g37192; PIDN:CAA45939.1; PDB:1ACcession: S20897

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 16330-16382,'S',16384-16756,'F',16758-16860 <LAB3>  
A:Cross-references: UNIPARC:UPI000016B0E6; EMBL:X64699; NID:g37190; PIDN:CAA45940.1; PDB:1

[illegible]







Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetellegre, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A./Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A./Reference number: A69580; MUID:98044033; PMID:9384377  
A./Accession: B69596

A./Status: nucleic acid sequence not shown; translation not shown

A./Molecule type: DNA

A./Residues: 1-1433 <KUN>

A./Cross-references: UNIPARC:UPI0000060385; GB:Z99111; GB:Z99112; GB:AL009126; NID:g26339

A./Experimental source: strain 168

C/Genetics:

A./Gene: bpr; bpf

A./Map position: 135 (degrees)

C./Superfamily: bacillopeptidase F; subtilisin homology

C./Keywords: extracellular protein; hydrolase; serine proteinase

F./1-30/Domain: signal sequence #status predicted <SIG>

F./31-194/Domain: propeptide #status predicted <PRO>

F./195-1433/Product: bacillopeptidase F #status experimental <MAT>

F./218-466/Domain: subtilisin homology <SBT>

F./227,274,452/Active site: Asp, His, Ser #status predicted

	Query Match	7.3%;	Score 89;	DB 1;	Length 1433;	
	Best Local Similarity	24.0%;	Pred. No. 21;			
	Matches	56;	Conservative	26;	Mismatches	71; Indels 80; Gaps 14
Qy	17 RMEGTVNGHEFEIEGEGEGRPEYEGHNTVTKLVTKGGLPFA-----					57
Db	521 KAEQVS-----VEGDDEPPVYQHKEV-TEAYEGSLPLTLTAEDNVSVTSVKLSYKLD					574
Qy	58 ---WDILSPQFOYGSKVYVKHPADIPDK--KLSPPEGFKMERVMNFEDG-----VVT					106
Db	575 QGEWTEITAKRISGDHLKGTYYAEIPIDKTJLS---YKW---MIHDFGHVSSDYVD					627
Qy	107 VTQDSSLQDGCFIYKKVF-----IGVN-----FPSDGPVMQKKTMGWEASTERLY					151
Db	628 VTVKPSITAG---YKQDFETAPGCWVASGTTNNMWEWGVPSTGP-----NTAASGEKVY					677
Qy	152 PRDGVLKGE-----IHKALKLKDGGHYLVFEKSIYMAKKPVQLPGYYYV					195
Db	678 GTN--LTGNYSANMANLVMPRIKAPDSGSLFLQFCKSWHNLEDDFDY-GYIVFV					727

RESULT 6  
B70476  
hypothetical protein aq\_2054 - Aquifex aeolicus  
C.Species: Aquifex aeolicus  
C.Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C.Accession: B70476  
R.Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998  
A.Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A.Reference number: A70300; MUID:98196666; PMID:9537320  
A.Accession: B70476  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-1116 <AQF>  
A.Cross-references: UNIPROT:O67838; UNIPARC:UP100000567A2; GB:AE000770; NID:g2984274; PI  
C.Experimental source: strain VF5  
C.Genetics:  
A.Gene: aq\_2054

Query Match	7.2%	Score 87.5;	DB 2;	Length 1116;
Best Local Similarity	22.0%;	Pred. No. 21;		
Matches	56;	Conservative 35;	Mismatches 103;	Indels 61;
			Gaps	9;
Qy	5	KNVKEFMRFKVRMEGTIVNGHFEIEGE	-----	GRPYEGHNTVRLKVTKGG 52
	:::	:::	:::	:::
Rb	164	KDILLDSEYFSAKTKGEIKRNTGEILAEVEIKERKENFTLISGTRKINAKGTINLPVLIDIN		223

[illegible]

```

RESULT 7
T06489
Probable peptidylprolyl isomerase (EC 5.2.1.8) FKBP77 - wheat
C/Species: Triticum aestivum (common wheat)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T06489
R/Brieman, A.
submitted to the EMBL Data Library, August 1996
A/Reference number: Z15713
A/Accession: T06489
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-568 <BRI>
A/Cross-references: UNIPROT:O04843; UNIPARC:UPI00000AD075; EMBL:Y07636; PIDN:CAA68913.1
A/Experimental source: cv. ATIR, 2 day old plants, root tips
C/Genetics:
A/Gene: FKBP77
C/Superfamily: peptidylprolyl isomerase ROF1; BKBP-type peptidylprolyl isomerase homology
C/Keywords: calmodulin binding; cis-trans-isomerase; cyclosporin A binding
F:54-101/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>
F:169-212/Domain: BKBP-type peptidylprolyl isomerase homology <PPI2>
F:286-334/Domain: BKBP-type peptidylprolyl isomerase homology <PPI3>
F:482-515/Domain: tetratricopeptide repeat homology <TTR>

```

	Query Match	7.2%	Score 87;	DB 2;	Length 568;
	Best Local Similarity	21.6%;	Pred. No. 9.3;		
	Matches	61;	Conservative	32;	Mismatches 81; Indels 108; Gaps 15
QY	5 KNVIKE-----FMRFKVRMEGT-VNGHFEIEEGEGRPYEGHNTVKKLVTKGGPL	54			
Db	39 KKLVEGEGGWDTAETALKVAVHTTGLTDGTFD-SSRDRGTFP-----KFLEQGQVI	91			
QY	55 PFAMD-----ILSPFOYGSKVTVKHPADIPDYKLSFP-----	88			
Db	92 K-GWDGIGIKTMKKGENASLTTPDLAYGERA---PRTIIPNATLPFDVELLSWASVKDI	146			
QY	89 -----EGFKWERVMNFEDGVVTVTQDSLDGCFTYKVFIGNVPSPDGPM	136			
Db	147 CKDGIFFKVLVEGQKWE---NPKDLDENVKYEARLEDGSVVSKESEI-----	192			
QY	137 QKKTGMWASTERLYPRDGLVKGEIHKALK-LKDGGHYLVEFSITY---MAKKPVQLPG	191			
Db	193 -----EFSV-----KDGIFCPALSAVKVTMKKGKVVLLTVKPQYGFGEQGRATEVEG	240			
QY	192 YYYVDSKLDI-----TSHNEDYTIV-----EQYER	216			
Db	241 AVPPNSTLIHIDLQLVSWKTLITLIGDDKRILKVLKEGEGYER	282			

RESULT 8  
C97196  
probable membrane protein [imported] - Clostridium acetobutylicum  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C/Accession: C97196







Db 406 VKTRKGIQTQLPFORRRPTRPLEVGAPFYQDGQLVKVYKTKTEDEPTVNRHYHVRWPEEA 465  
QY 91 FKWERVWNFEDGCVTVTQDSSLQDGCFFIYKVFIGNFPPSDGPMQKTMGEASTERL 150  
Db 466 CAHNRRTGSEASSGMTHENYIILQDLSFSCKYK-----TVQPIRPKS---HSKAEAV 515  
QY 151 Y---PRDGVLKGEIHKALK-LKDGGHYLVEFKSIYMAKKPVQLPGYYVVDSKLDITSH 204  
Db 516 FFTTPPCSALKGSKHKPIGCLGEGHYLSK-----VLAKPENLSASFIVQD-VNITGH 567

RESULT 12

IUMSNG  
neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse

N;Alternate names: NCAM-120

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004

C;Accession: A29673; S00382; A44290

R;Bartheleis, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaux, J.C.; Hirsch, M.R.; Fontec

EMBO J. 6, 907-914, 1987

A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000

A;Reference number: A29673; MUID:87246524; PMID:3595563

A;Accession: A29673

A;Molecule type: mRNA

A;Residues: 1-725 <BAR>

A;Cross-references: UNIPROT:P13594, UNIPARC:UPI0000028AC3, EMBL:Y00051, NID:G53342; PIDN

R;Barbas, J.A.; Chaux, J.C.; Steinmetz, M.; Goridis, C.

EMBO J. 7, 625-632, 1988

A;Title: Differential splicing and alternative polyadenylation generates distinct NCAM v

A;Reference number: S00382; MUID:88283628; PMID:3396534

A;Accession: S00382

A;Molecule type: DNA

A;Residues: 642-656, 'D', 658-725 <BA2>

A;Cross-references: UNIPARC:UPI0000174377; EMBL:X07195

R;Rougon, G.; Marhak, D.R.

J. Biol. Chem. 261, 3396-3401, 1986

A;Title: Structural and immunological characterization of the amino-terminal domain of m

A;Reference number: A44290; MUID:86140120; PMID:3512556

A;Accession: A44290

A;Molecule type: protein

A;Residues: 20-36 <ROU>

A;Cross-references: UNIPARC:UPI0000174376

C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol

C;Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IUMS

C;Genetics:

A;Gene: NCAM

A;Map position: 9

A;introns: 701/1

C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

C;Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane

F;1-19/Domain: signal sequence #status predicted <SIG>

F;34-98/Domain: immunoglobulin homology <IMM1>

F;132-191/Domain: immunoglobulin homology <IMM2>

F;152-156/Region: heparin binding #status predicted

F;161-165/Region: heparin binding #status predicted

F;228-290/Domain: immunoglobulin homology <IMM3>

F;263-272/Region: NCAM binding #status predicted

F;323-388/Domain: immunoglobulin homology <IMM4>

F;420-482/Domain: immunoglobulin homology <IMM5>

F;519-596/Domain: fibronectin type III repeat homology <FN3A>

F;625-685/Domain: fibronectin type III repeat homology <FN3B>

F;41-96,139-189,235-288,330-386,427-480/Dsulfide bonds: #status predicted

F;222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.0%; Score 84.5; DB 1; Length 725;  
Best local similarity 24.2%; Pred. No. 21;  
Matches 37; Conservative 30; Mismatches 57; Indels 29; Gaps 9;

QY 91 FKWERVWNFEDGCVTV--TQDSSLQDGCFFIYKVFIGNFPPSDGPMQKTMGEAST 147  
Db 548 FTWYDAKEANNEGIVTIMGLKPETYS-----RLAALNGKLGELIMQPSSESKTQVPV 600  
QY 148 ERLYPRDGVLKGEIHK-----ALKLKDGG---HYLVEFKSIYMAKKP-VQLP-GYY 193

Db 601 ELSAPK--LEGQMGEDGNSIKVNLIKQDDGSPIRHYLVKTRALASEWKPRIPLSGSH 657  
QY 194 YVDSK-LDITSHNEDYTIIVE-QYERTEGRHHLF 224  
Db 658 HVMKSLDWNABEYEVYVAENQGGKSKAAHFVF 690

RESULT 13

T17884

S-layer protein - Bacillus circulans

C;Species: Bacillus circulans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T17884

R;Aubert-Pivert, E.; Davies, J.

Gene 147, 1-11, 1994

A;Title: Biosynthesis of butirosin in Bacillus circulans NRRL B3312: identification by se

A;Reference number: Z18808; MUID:94374689; PMID:7522196

A;Accession: T17884

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1616 <AUB>

A;Cross-references: UNIPROT:P35824, UNIPARC:UPI0000135A00; EMBL:L20421; NID:G304142; PID:

C;Genetics:

A;Gene: butB

A;Function: butirosin biosynthesis

Query Match 6.9%; Score 84; DB 2; Length 1616;  
Best local similarity 20.6%; Pred. No. 69;  
Matches 59; Conservative 48; Mismatches 85; Indels 94; Gaps 17;

QY 6 NVIKEMR-----FKVMEGTVNGHEFEI--EGEGGRPYEGHNTVKLKVTX 50  
Db 590 DLQEFIRYSRELGLDIHVSFNIFAEGSIASNEFALDSDHLDWEERVYNAADNGQIK--- 646  
QY 51 GGPLPFAWDILSPQFGYSKYVVKHPAD-IPDYKSLSPESGFKMERVMNFEDGCV--TV 107  
Db 647 -----RLRESAKQGAVALFVNPNSNDEVDPQLKTI-----EVLQNYDVQGVLDRA 692  
QY 108 TQDSSLQDGCFFIYKVK--FIGV-----NFP-----SDGPMQ-----KKT 140  
Db 693 RYDMSADFSDLTKAKFESFLGARQKQLQWPPDVFYAGNVKRDGPLIRDWBFRSKTI 752  
QY 141 MGWEASTERLYPRDGVLKGEIHKALKLK-----DGHY-LVEFK--SIYMAK 184  
Db 753 KSFTEVVRQLTDR---VKAEKQKIEVSAYVGSWFESYYLVGVHWGSTEFPRYDERLRMKD 809  
QY 185 KPVQLPGYY--YVDSKLDITSHNEDYTIIVEQYERT--EGRHHLFL 225  
Db 810 KSVYTPGYYESGV-----KNULDIFIMIGAYQTAPAEIHYITL 847

RESULT 14

E90276

conserved hypothetical protein [imported] - Sulfolobus solfataricus

C;Species: Sulfolobus solfataricus

C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C;Accession: E90276

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Y

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A;Description: Sulfolobus solfataricus complete genome.

A;Reference number: A99139

A;Accession: E90276

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-268 <KUR>

A;Cross-references: UNIPROT:Q97YT6, UNIPARC:UPI00000643B4; GB:AE006641; NID:G13814420; PJ

C;Genetics:

A;Gene: SSO1221



Query Match	6.9%;	Score 83.5;	DB 2;	Length 268;
Best Local Similarity	25.8%;	Pred. No. 7;		
Matches	49;	Conservative	21;	Mismatches 59; Indels 61; Gaps 13;
OY	34	EGRPYEGHNTVTKLVTKG-----PLPFAWDILSPQFGYSKY-----VKHPAD	78	
DB	3	EIRPY--YNTFSLPITKSGKSQIVPPPPWYAI---EMIGKAYFDPVKVLDLVPPPLE	56	
OY	79	IPD-----YKLSFPEGFKMERVMPEDGCVTVTQDSSLQDGCFTYVKF---IGVNF	129	
DB	57	IYDGEGFYIAKIFTVSGNRWE--MLYED-----PEETKYMAAIALKVKYNDNIFTYF	108	
OY	130	P-----SDGPMQKKTMGWEASTERLYPRD--GVLKGEIHKALK-----LKDGHYL	174	
DB	109	PFMMVVDKDLPLR---GWLIG---YPKGLAYISISFHKLLDGYSGPSSGVRMGGYAL	160	
OY	175	VEFKSIYMAK	184	
DB	161	RNGKEIRVK	170	

RESULT 15  
 E69544  
 hypothetical protein AF2357 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: E69544  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997  
 A:Authors: Uterback, T.; Colton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: E69544  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-340 <KLE>  
 A:Cross-references: UNIPROT:O30313; UNIPARC:UPI00000572DC; GB:AE001113; GB:AE000782; NID

Query Match	6.9%;	Score 83.5;	DB 2;	Length 340;
Best Local Similarity	23.9%;	Pred. No. 9.6;		
Matches	38;	Conservative	29;	Mismatches 63; Indels 29; Gaps 9;
OY	41	HNTVKLVTKGGLPFAWDILSPQFGYSKYVKGHPADIPDYKLSFPEGFKMERVMP	100	
DB	144	HAWVEVKINN-----SWVADPTVYV--LYVNYPEKYPNMNKLWF--NNESWANLIDF-	192	
OY	101	DGGVTVTQDSSLQDGCFTYVKVFIQVNFPSDGPVQ--KKTMGWEASTERLYPRDGVLK	158	
DB	193	-SRVTVLPNGSVLDLTSNY-TKTYNVTITIDQNVKGIKVTTWKGSVERTVYSKAVNK	250	
OY	159	GE-IHKALKDGHVLEFKSIYMAKPVQLPGYYVD	196	
DB	251	SDTVNLALATR-----IYFELI-----VPTWYFLE	276	

Search completed: January 11, 2006, 02:04:50  
 Job time : 18 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 01:54:16 ; Search time 73 Seconds  
(without alignment)  
2174.574 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHLLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	*Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	1 RFP DISSP	Q9u6y8 discosoma s
2	1196	98.5	236	2 Q66ND7_9CNID	Q66nd7 discosoma s
3	1193	98.3	225	2 Q6KF85_9CNID	Q6kf85 discosoma s
4	1191	98.1	236	2 Q66ND8_9CNID	Q66nd8 discosoma s
5	1085.5	89.4	230	2 Q9GTU7_9CNID	Q9gtj7 discosoma s
6	985	81.1	226	2 Q5S3G8_9CNID	Q5s3g8 discosoma s
7	972	80.1	226	2 Q5S3G7_9CNID	Q5s3g7 discosoma s
8	804	66.2	221	2 Q66PV1_ACRTE	Q66pv1 acropora te
9	798	65.7	221	2 Q66PV8_9CNID	Q66pv8 acropora ac
10	797	65.7	221	2 Q66PV0_ACRMT	Q66pv0 acropora mi
11	793	65.3	221	1 NFPC GONTE	Q95p04 gonlopora t
12	793	65.3	221	2 Q66PU9_9CNID	Q66pu9 acropora hy
13	782	64.4	221	1 NFPC MONER	P83690 montipora e
14	745	61.4	227	2 Q66ND6_DISSP	Q66nd6 discosoma s
15	729.5	60.1	232	1 GFPL DISSP	Q9u6y7 discosoma s
16	690	56.8	225	2 Q963F5_9CNID	Q963f5 montastraea
17	684	56.3	225	2 Q66ND3_9CNID	Q66nd3 montastraea
18	681	56.1	225	2 Q720W4_9CNID	Q720w4 montastraea
19	681	56.1	225	2 Q95UA7_9CNID	Q95ua7 montastraea
20	679.5	56.0	227	2 Q720W6_9CNID	Q720w6 montastraea
21	678.5	55.9	227	2 Q720W8_9CNID	Q720w8 montastraea
22	678.5	55.9	227	2 Q962P9_9CNID	Q962p9 montastraea
23	678	55.8	225	2 Q720W5_9CNID	Q720w5 montastraea
24	677	55.8	224	2 Q5TUG6_9CNID	Q5tug6 echinophyll
25	672.5	55.4	227	2 Q66ND2_9CNID	Q66nd2 montastraea
26	670.5	55.2	227	2 Q66ND5_9CNID	Q66nd5 montastraea
27	669.5	55.1	234	2 Q720W7_9CNID	Q720w7 montastraea
28	668.5	55.1	234	2 Q8TSF2_9CNID	Q8tsf2 montastraea
29	667	54.9	225	2 Q60125_9CNID	Q60125 galaxea fas
30	667	54.9	266	1 GFPL CIASP	Q9u6y3 clavularia
31	663	54.6	226	2 Q8T6U0_9CNID	Q8t6u0 dendronephc

32	660.5	54.4	227	2 Q66ND4_9CNID	Q66nd4 montastraea
33	655.5	54.0	226	2 Q66PU5_9CNID	Q66pu5 agaricia fr
34	655.5	54.0	227	2 Q95VT0_9CNID	Q95vt0 montastraea
35	654.5	53.9	234	2 Q8MU47_9CNID	Q8mu47 montastraea
36	652.5	53.7	225	2 Q720W9_9CNID	Q720w9 montastraea
37	649.5	53.5	227	2 Q53UG7_9CNID	Q53ug7 favia favus
38	649	53.5	225	2 Q60124_9CNID	Q60124 galaxea fas
39	642.5	52.9	225	2 Q6USK3_9CNID	Q6usk3 montastraea
40	639.5	52.7	227	2 Q53UG8_9CNID	Q53ug8 favia favus
41	639	52.6	224	2 Q8MU48_9CNID	Q8mu48 montastraea
42	631.5	52.0	225	2 Q8TSF1_9CNID	Q8tsf1 montastraea
43	620.5	51.1	230	2 Q66PW1_9CNID	Q66pw1 scolymia cu
44	610	50.2	223	2 Q6R8F5_9CNID	Q6r8f5 astrangia l
45	608.5	50.1	231	1 RFP_PARC	Q8isf8 parasicyoni

ALIGNMENTS

RESULT 1  
RFP DISSP STANDARD; PRT; 225 AA.  
ID RFP DISSP  
AC Q9u6y8;  
DT 01-FEB-2005 (Rel. 46, Created)  
DT 01-FEB-2005 (Rel. 46, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Red fluorescent protein dRFP583 (DRed).  
OS Discosoma sp. (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=86600;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=99436614; PubMed=10504696; DOI=10.1038/13657;  
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,  
RA Markelov M.L., Lukyanov S.A.;  
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";  
RL Nat. Biotechnol. 17:969-973(1999).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX PubMed=1101896; DOI=10.1038/81992;  
RA Wall M.A., Socolich M., Ranganathan R.;  
RT "The structural basis for red fluorescence in the tetrameric GFP  
homolog DRed.";  
RL Nat. Struct. Biol. 7:1133-1138(2000).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND SUBUNIT.  
RX PubMed=11209050; DOI=10.1073/pnas.98.2.462;  
RA Yarbrough D., Wachter R.M., Kalilo K., Matz M.V., Remington S.J.;  
RT "Refined crystal structure of DRed, a red fluorescent protein from  
coral, at 2.0-A resolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:462-467(2001).  
CC -I- FUNCTION: Thought to play a role in photoprotection of the coral's  
resident symbiont microalgae's photosystems from photoinhibition  
caused by high light levels found near the surface of coral reefs.  
CC In deeper water, the fluorescence may be to convert blue light  
into longer wavelengths more suitable for use in photosynthesis by  
the microalgal symbionts.  
CC -I- BIOPHYSICOCHEMICAL PROPERTIES:  
CC Absorption:  
CC Abs(max)=558 nm;  
CC Note=Exhibits a smaller absorbance peak at 494 nm. The broad  
fluorescence emission spectrum peaks at 583 nm;  
CC -I- SUBUNIT: Homotetramer.  
CC -I- PTM: Contains a chromophore consisting of modified amino acid  
residues. The chromophore is formed by autocatalytic backbone  
condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1)  
to dihydroxyrosine, and formation of a double bond to the alpha-  
amino nitrogen of residue Xaa-N. Maturation of the chromophore  
requires nothing other than molecular oxygen.  
CC -I- BIOTECHNOLOGY: Fluorescent proteins have become a useful and  
ubiquitous tool for making chimeric proteins, where they function  
as a fluorescent protein tag. Typically they tolerate N- and C-



```
CC      terminal fusion to a broad variety of proteins. They have been
CC      expressed in most known cell types and are used as a noninvasive
CC      fluorescent marker in living cells and organisms. They enable a
CC      wide range of applications where they have functioned as a cell
CC      lineage tracer, reporter of gene expression, or as a measure of
CC      protein-protein interactions.
CC      -I- SIMILARITY: Belongs to the GFP family.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
DR      EMBL; AF168419; AAF03369.1; -; mRNA.
DR      PDB; 1G7K; X-ray; A/B/C/D=2-225.
DR      PDB; 1GGX; X-ray; A/B/C/D=1-225.
DR      InterPro; IPR01584; GFP_related.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFP.
DR      ProDom; PD013756; Green_fl_protein; 1.
KW      3D-structure; Chromophore; Luminescence; Photoprotein.
FT      MOD_R13      67      (Z)-2,3-didehydrotyrosine.
FT      CROSSTANK      66      68      2-iminomethyl-5-imidazolinone (Gln-Gly).
SQ      SEQUENCE      225 AA; 25931 MW; FBFA5369778F689 CRC64;

Query Match      100.0%; Score 1214; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 1e-96;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
        |||
        1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
        |||
        61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||
        121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
DB      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
        181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
DB      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 2
Q66ND7_9CNID
ID      Q66ND7_9CNID PRELIMINARY; PRT; 236 AA.
AC      Q66ND7;
DT      25-OCT-2004 (Tremblrel. 28, Created)
DT      25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT      25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE      Enhanced red fluorescent protein R+.
OS      Discosoma sp. RC-2004.
OC      Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC      Discosomatidae; Discosoma.
OX      NCBI_TaxID=289055;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT      "Cloning of Anthozoa Fluorescent Protein Genes.";
RL      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY679107; AAU04444.1; -; mRNA.
DR      GO; GO:0006091; P:generation of precursor metabolites and energy; IBA.
DR      InterPro; IPR01584; GFP_related.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFP.
DR      ProDom; PD013756; Green_fl_protein; 1.
SQ      SEQUENCE      236 AA; 27032 MW; BB3844BE06829EF0 CRC64;
```

```
Query Match      98.5%; Score 1196; DB 2; Length 236;
Best Local Similarity 98.2%; Pred. No. 3.9e-95;
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
        |||
        1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
        |||
        61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||
        121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
DB      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
        181 YMAKKPVQLPGYYVVDKLDITSHNKDYTIIVEQYERTEGRHHLFL 225
DB      181 YMAKKPVQLPGYYVVDKLDITSHNKDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 3
Q6KF85_9CNID
ID      Q6KF85_9CNID PRELIMINARY; PRT; 225 AA.
AC      Q6KF85;
DT      05-JUL-2004 (Tremblrel. 27, Created)
DT      05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT      05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE      Orange fluorescent protein FP586.
OS      Discosoma sp. JW-2002.
OC      Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC      Discosomatidae; Discosoma.
OX      NCBI_TaxID=208461;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Wiedenmann J., Girod A.;
RL      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF545828; AAQ11987.1; -; mRNA.
DR      HSSP; P42212; 1B9C.
DR      GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR      InterPro; IPR01584; GFP_related.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFP.
DR      ProDom; PD013756; Green_fl_protein; 1.
SQ      SEQUENCE      225 AA; 25791 MW; E151D0E497AA23FA CRC64;
```

```
Query Match      98.3%; Score 1193; DB 2; Length 225;
Best Local Similarity 98.2%; Pred. No. 6.7e-95;
Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
        |||
        1 MSCSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMD 60
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
        |||
        61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTY 120
QY      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||
        121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
DB      121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||
        181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
DB      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 4
Q66ND8_9CNID
ID      Q66ND8_9CNID PRELIMINARY; PRT; 236 AA.
AC      Q66ND8;
```







AC Q5S3G7;  
DT 01-FEB-2005 (TREMBLrel. 29, Created)  
DT 01-FEB-2005 (TREMBLrel. 29, last sequence update)  
DT 01-FEB-2005 (TREMBLrel. 29, last annotation update)  
DE Fluorescent protein plun.  
OS Discosoma sp. LW-2004.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=301246;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15556995; DOI=10.1073/pnas.0407752101;  
RA Wang L., Jackson W.C., Steindach P.A., Tsien R.Y.;  
RT "Evolution of new nonantibody proteins via iterative somatic  
RT hypermutation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
DR EMBL; AY786537; AAV65487.1; -; mRNA.  
DR SMR; Q5S3G7; 8-222.  
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
DR InterPro; IPR011584; GFP related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFPUNESCENT.  
DR Prodom; PD013756; Green\_fl\_protein; 1.  
SQ SEQUENCE 226 AA; 25590 MW; 6DF38CAD2AB28BED CRC64;

Query Match	80.1%	Score 972;	DB 2;	Length 226;
Best Local Similarity	83.4%	Pred. No. 8.4e-76;		
Matches 181; Conservative	8;	Mismatches 28;	Indels 0;	Gaps 0;

```

QY      5 KQV|KEFMRFKVRMEGTUNGHEFEIEGEGEGRPYEGHNTV|KLVTKGGRPLPFAMDI|LSPQ 64
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db      6 EVV|KEFMRFKEHMEGSVNGHEFEIEGEGEGRPYEGTQTAR|LKVTKGGPLPFAMDI|LSPQ 65

QY      65 FQYGSKVYV|KHPADI|PDYK|LSP|EGGFKWERV|NPFEDG|VTVTQDSS|LQDGC|FYKV|KF 124
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      66 IMYGSKAYV|KHPADI|PDY|L|K|LSP|EGGFKWERV|NPFEDG|VTVTQDSS|LQDGC|FYKV|KF 125

QY      125 IGVNPPSDGPV|MQKKTGMWEAST|RLYP|RDGYL|KGEI|H|KAL|KL|DGGHYL|VEFKSI|YMAK 184
      | ||||| ||||| ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db      126 RGVNPPSDGPV|MQKKTGMWEAST|RLYP|RDGYL|KGEI|H|KAL|KL|DGGHYL|VEFKSI|YMAK 185

QY      185 KPVQDLPGY|YVDSK|LDITSHNEDY|TIVEQYERT|EGRH 221
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      186 KPVQDLPGY|YVDSK|LDITSHNEDY|TIVEQYERT|EGRH 222

```

## RESULT 8

Q66PV1 ACRTE  
ID Q66PV1 ACRTE PRELIMINARY; PRT; 221 AA.

NC 25-OCT-2004 (TReMBLrel. 28, Created)  
 DT 25-OCT-2004 (TReMBLrel. 28, last sequence update)  
 DT 25-OCT-2004 (TReMBLrel. 28, last annotation update)  
 DE Chromoprotein.  
 OS Acropora tenuis (Purple tipped acropora).  
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
 OC Asterozoenina; Acroporidae; Acropora.  
 OX NCBI\_TaxID=70783;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;  
 RT "Survey of coral GFP-like proteins."  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY646074; AAU06853.1; -; mRNA.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA  
 DR InterPro; IPR011584; GFP\_related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP\_1.  
 DR PRINTS; PRO1229; GFP\_LUORESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 221 AA; 25103 MW; 3BEA8B848C8F5DA4 CRC64;

Query Match	66.2%;	Score 804;	DB 2;	Length 221;
Best Local Similarity	67.9%;	Pred. No. 2.7e-61;		
Matches 148;	Conservative 25;	Mismatches 45;	Indels 0;	Gaps 0;

Oy	6	NVIKEFMRFKVRMEGTIVNGHEHFEIEGEGEGRPYEGHNTVKLKVTKGGPLPFAMDILSPQF	65
Dp	2	SVAIAKQMTYKVMYSGTIVNGHYFEVEGDGKKPYEGEQTVKLVTIKGGPLPFAMDILSPQF	61
Oy	66	QYGSKVYVKHPADIPDYKLTSPPEGFKMERVMNPFEDGGVWTVTODSLDCGFYIKVKPI	125
Dp	62	QYGSIFPTKYPEDIPDYVXQSFPBGYTWERIMNFEDDACTVSNDSSIQCNCFYHVKS	121
Oy	126	GWNFPSDGPVMQOKKTMGWEASTERLYPRDGVLKGBIHKALKLKGCGHYLVYEKSIYMAKK	185
Dp	122	GWNFPNPNGPVMQOKKTQGWEPEINTERLLFARDGMILIGNFMALKEGGGHYLCEKSTYKAkk	181
Oy	186	PVQLPGYYVVDskLDITSHNEDYTIIVEQYERTEGRHHL	223
Dp	182	PVRMPGYHYVDRKLDVTNNHNKDYTSVEQCEISIAARKPL	219

## RESULT 9

Q66PU8 9CNID  
ID Q66PU8 9CNID PRELIMINARY; PRT; 221 AA.

DT 25-OCT-2004 (TREMBLrel. 28, Created)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE Chromoprotein.  
 OS Acropora aculeus.  
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
 OC Astrocenina; Acroporidae; Acropora.  
 OX NCBI\_TaxID=287157;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Alieva N.O., Meleshkevitch E.A., Field S.F., Matz M.V.;  
 RT "Survey of coral GFP-like proteins."  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY646077; AAU06856.1; -; mRNA.  
 DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA  
 DR InterPro; IPR011584; GFP related.  
 DR InterPro; IPR000786; Green\_fl\_protein.  
 DR Pfam; PF01353; GFP; 1.  
 DR PRINTS; PR01229; GFPUNRESCENT.  
 DR ProDom; PD013756; Green\_fl\_protein; 1.  
 SQ SEQUENCE 221 AA; 24976 MW; 307DE4FC4B018B43 CRC64;

Query Match 65.7%; Score 798; DB 2; Length 221;

Best Local Similarity 69.0%; Pred. No. 8.9e-61;  
Matches 145; Conservative 24; Mismatches 41; Indels 0; Gaps 0;

[illegible]

## RESULT 10

Q66PV0_ACRMI	
ID Q66PV0_ACRMI	PRELIMINARY;
AC Q66PV0;	PRT;
DT 25-OCT-2004	221 AA.
(TREMBlrel. 28, Created)	







```

DR EMBL; AY646076; AAU06855.1; -; mRNA.
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
DR SQQ SEQUENCE 221 AA; 24963 MW; FC9D9E749069D079 CRC64;

Query Match          65.3%; Score 793; DB 2; Length 221;
Best Local Similarity 68.6%; Pred. No. 2.4e-60;
Matches 144; Conservative 25; Mismatches 41; Indels 0; Gaps 0

QY 6 NVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPRYEGHNTVKLKVTKGGLPFAMDLSPQF 65
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 2 SVIATQMTYKVMMSGTVNGHYFEVEGDKKPRYEGEGTVRLTVTKGGLPFAMDLSPQS 61

QY 66 QYGSKVYVKHPADIPDYKKLSPEEGFKMERVMNFEDGCVTVTQDSSLQDGCPIYKKEI 125
   |||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 62 QYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQNCFIYHKFS 121

QY 126 GVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKGELHKAALKDGGHYLVFKSIYMAK 185
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 122 GLNFPNGPVMQKKTQGWEPNTERLFARDGVLIGNNFMAKLKLEGGHYLCFESKITYAKK 181

QY 186 PVQLPGYYVDSKLDITSHNEDYTIIVEQYE 215
   ||:||||:|||||:||||:|||||
Db 182 PVKMPGYHFVDRKLDVTNHNKDYTSVEQRE 211

RESULT 13
NFCP_MONEF STANDARD; PRT; 221 AA.
ID NFCP_MONEF
AC P83630;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE GFP-like non-fluorescent chromoprotein (Rtms 5) (Non-fluorescent
DE pocilloporin).
DE Montipora efflorescens (Coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Asterozoenina; Acroporidae; Montipora.
OX NCBI_TaxID=105610;
RN [1]
RP CRYSTALLIZATION.
RX PubMed=12595737; DOI=10.1107/S0907444902023466;
RA Beddoe T., Ling M., Dove S., Hoegh-Guldberg O., Devenish R.J.,
RA Prescott M., Rossjohn J.;
RT "The production, purification and crystallization of a pocilloporin
RT pigment from a reef-forming coral.";
RL Acta Crystallogr. D 59:597-599(2003).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS), SUBUNIT, AND MUTAGENESIS OF
RP HIS-142 AND PHE-158.
RX MEDLINE=22511991; PubMed=12623015; DOI=10.1016/S0969-2126(03)00028-5;
RA Prescott M., Ling M., Beddoe T., Oakley A.J., Dove S.,
RA Hoegh-Guldberg O., Devenish R.J., Rossjohn J.;
RT "The 2.2 A crystal structure of a pocilloporin pigment reveals a
RT nonplanar chromophore conformation.";
RL Structure 11:275-284(2003).
CC -1- FUNCTION: Thought to play a role in photoprotection of the coral's
CC resident symbiont microalgae's photosystems from photoinhibition
CC caused by high light levels found near the surface of coral reefs.
CC -1- SUBUNIT: Homotetramer.
CC -1- PTM: Contains a chromophore consisting of modified amino acid
CC residues. The chromophore is formed by autocatalytic backbone
CC condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1)
CC to dihydroxytyrosine, and formation of a double bond to the alpha-
CC amino nitrogen of residue Xaa-N. Maturation of the chromophore
CC requires nothing other than molecular oxygen.
CC -1- BIOTECHNOLOGY: Fluorescent proteins have become a useful and
CC ubiquitous tool for making chimeric proteins, where they function
CC as a fluorescent protein tag. Typically they tolerate N- and C-
CC terminal fusion to a broad variety of proteins. They have been

```

```

CC expressed in most known cell types and are used as a noninvasive
CC fluorescent marker in living cells and organisms. They enable a
CC wide range of applications where they have functioned as a cell
CC lineage tracer, reporter of gene expression, or as a measure of
CC protein-protein interactions.
CC -1- MISCELLANEOUS: The wild-type form is non-fluorescent with the
CC color being pH dependent, ranging from yellow at low pH, through
CC red to blue at high pH. The His-142 mutation produces a
CC fluorescent form.
CC -1- SIMILARITY: Belongs to the GFP family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PDB; 1MOU; X-ray; A=3-221.
CC PDB; 1MOV; X-ray; A=3-221.
CC DR GO; GO:0018995; C:host; TAS.
CC DR GO; GO:0010117; P:photoprotection; TAS.
CC DR InterPro; IPR011584; GFP_related.
CC DR InterPro; IPR000786; Green_fl_protein.
CC DR Pfam; PF01353; GFP; 1.
CC DR ProDom; PD013756; Green_fl_protein; 1.
CC KM 3D-structure; Chromophore; Luminescence; Photoprotein.
CC FT MOD_RES 63 63 (E)-2,3-didehydrotyrosine.
CC FT CROSSLINK 62 64 2-iminomethyl-5-imidazolnolone (Gln-Gly) .
CC FT MUTAGEN 142 142 H->S: Produces a fluorescent form.
CC FT MUTAGEN 158 158 F->H: Produces a homodimeric form.
CC SQ SEQUENCE 221 AA; 24911 MW; 626B1A7BEDD7393B CRC64;

Query Match 64.4%; Score 782; DB 1; Length 221;
Best Local Similarity 68.1%; Pred. No. 2.1e-59;
Matches 143; Conservative 23; Mismatches 44; Indels 0; Gaps 0;

QY 6 NVIKEFMRFKRMEGTVNGHFEIEGEGEGRPYEGHNTVTKLVTCKGRLPFAMDILSPQF 65
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 2 SVIATQMTYKVVMSGTVNGHYFEVEGDKGRPYEGEQTKLTVTKGGPLPFAMDILSPQC 61

QY 66 QYGSKVVYKHPADIPDYKLLSPPEGFKERVMNPFEDGGVTVTQDSSLQDGGRIYKVKFI 125
   |||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 62 QYGSIPFTKYPEDIPDYVKQSFPEGFTWERIMNFEDGAVCTVSNDSSIQNCFTYHVKS 121

QY 126 GVNFPSDGPVMOCKTMGWEASTERLPRDGLVKGELHKLKLDKGHYLVEFKSIYMAKK 185
   |||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 122 GLNFPNPGVPMQKKTQGWEPHSERLFAFGCMLIGNFMALKLEGGHYLCFEKTTYKAKK 181

QY 186 PVQLPGYVVYVDSKLDITSHNEDYTIIVEQYE 215
   |||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 182 PVKMPGYHYVDRLDVTNHNKDYTSVEQCE 211

RESULT 14
Q66ND6_DISST
ID Q66ND6_DISST PRELIMINARY; PRT; 227 AA.
AC Q66ND6;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Cyan fluorescent protein C1.
OS Discosoma striata (Striped mushroom).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=105400;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT "Cloning of Cnidarian Fluorescent Protein Genes.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY679108; AAU04445.1; -; mRNA.
DR GO; GO:006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP_related.

```







**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using BW model

Run on: January 11, 2006, 01:59:57 ; Search time 23 Seconds  
(without alignments)  
808.784 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1214	100.0	225	2	US-09-866-538-12	Sequence 12, Appl
2	1214	100.0	225	2	US-09-865-291-12	Sequence 12, Appl
3	1210	99.7	240	2	US-10-152-296-2	Sequence 2, Appl
4	681	56.1	225	2	US-10-244-779-2	Sequence 2, Appl
5	485.5	40.0	238	2	US-09-277-716-16	Sequence 16, Appl
6	485.5	40.0	238	2	US-09-609-161B-16	Sequence 16, Appl
7	485.5	40.0	238	2	US-09-626-581D-65	Sequence 65, Appl
8	485.5	40.0	238	2	US-09-415-765B-65	Sequence 65, Appl
9	485.5	40.0	238	2	US-09-626-580C-65	Sequence 65, Appl
10	473.5	39.0	224	2	US-09-977-897-3	Sequence 3, Appl
11	473.5	39.0	224	2	US-09-977-897-13	Sequence 13, Appl
12	473.5	39.0	225	2	US-09-977-897-14	Sequence 14, Appl
13	473.5	39.0	226	2	US-09-977-897-15	Sequence 15, Appl
14	473.5	39.0	227	2	US-09-977-897-16	Sequence 16, Appl
15	473.5	39.0	228	2	US-09-977-897-17	Sequence 17, Appl
16	473.5	39.0	229	2	US-09-977-897-18	Sequence 18, Appl
17	473.5	39.0	230	2	US-09-977-897-19	Sequence 19, Appl
18	473.5	39.0	231	2	US-09-977-897-20	Sequence 20, Appl
19	473.5	39.0	232	2	US-09-977-897-21	Sequence 21, Appl
20	473.5	39.0	233	2	US-09-977-897-22	Sequence 22, Appl
21	473.5	39.0	234	2	US-09-977-897-9	Sequence 9, Appl
22	473.5	39.0	234	2	US-09-977-897-23	Sequence 23, Appl
23	473.5	39.0	235	2	US-09-977-897-8	Sequence 8, Appl
24	473.5	39.0	235	2	US-09-977-897-24	Sequence 24, Appl
25	473.5	39.0	236	2	US-09-977-897-7	Sequence 7, Appl
26	473.5	39.0	236	2	US-09-977-897-25	Sequence 25, Appl
27	473.5	39.0	237	2	US-09-977-897-6	Sequence 6, Appl

28	473.5	39.0	237	2	US-09-977-897-26	Sequence 26, Appl
29	473.5	39.0	238	2	US-09-277-716-32	Sequence 32, Appl
30	473.5	39.0	238	2	US-09-609-161B-32	Sequence 32, Appl
31	473.5	39.0	238	2	US-09-839-650-3	Sequence 3, Appl
32	473.5	39.0	238	2	US-09-977-897-5	Sequence 5, Appl
33	473.5	39.0	238	2	US-09-977-897-27	Sequence 27, Appl
34	473.5	39.0	238	2	US-10-652-703A-3	Sequence 3, Appl
35	473.5	39.0	239	2	US-09-977-897-2	Sequence 2, Appl
36	473	39.0	219	2	US-09-977-897-4	Sequence 4, Appl
37	473	39.0	231	2	US-09-977-897-12	Sequence 12, Appl
38	473	39.0	232	2	US-09-977-897-11	Sequence 11, Appl
39	473	39.0	233	2	US-09-977-897-10	Sequence 10, Appl
40	465.5	38.3	239	2	US-10-021-818A-4	Sequence 4, Appl
41	464.5	38.3	238	2	US-10-021-818A-2	Sequence 2, Appl
42	215.5	17.8	238	2	US-09-023-946B-23	Sequence 23, Appl
43	214.5	17.7	238	2	US-09-023-946B-28	Sequence 28, Appl
44	213.5	17.6	240	2	US-09-129-192C-49	Sequence 49, Appl
45	212.5	17.5	238	1	US-08-818-604-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1  
US-09-866-538-12  
; Sequence 12, Application US/09866538  
; Patent No. 6852849  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/09/866,538  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29.  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPYEGHNTVKLKVTGKGPLPFAWDI	60
DB	1	MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGRPYEGHNTVKLKVTGKGPLPFAWDI	60
QY	61	LSPOFQYGSKVYKHPADIPDYKLSFPEGFKWERVWNPEDGVTVTQDSSLQDGCFTY	120
DB	61	LSPOFQYGSKVYKHPADIPDYKLSFPEGFKWERVWNPEDGVTVTQDSSLQDGCFTY	120
QY	121	KVKEIGVNFPSDGPVMOCKTMGWEASTERLYPRDGLKGEIHKALKLKDGHYLVFKSI	180
DB	121	KVKEIGVNFPSDGPVMOCKTMGWEASTERLYPRDGLKGEIHKALKLKDGHYLVFKSI	180
QY	181	YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL	225
DB	181	YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL	225

RESULT 2  
US-09-865-291-12  
; Sequence 12, Application US/09865291  
; Patent No. 6900304  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: TING, Alice  
; APPLICANT: ZHANG, Jin



; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Diacosoma sp.
US-09-865-291-12

Query Match 100.0%; Score 1214; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 3e-132;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLTGKGGPLPFAWDI 60
|
DB 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLTGKGGPLPFAWDI 60
61 LSPQFGYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120
|
61 LSPQFGYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRRHHLFL 225
|
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRRHHLFL 225

RESULT 3
US-10-152-296-2
; Sequence 2, Application US/10152296
; Patent No. 6723537
; GENERAL INFORMATION:
; APPLICANT: Peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DSRED) of Diacosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
US-10-152-296-2

Query Match 99.7%; Score 1210; DB 2; Length 240;
Best Local Similarity 99.6%; Pred. No. 9.5e-132;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLTGKGGPLPFAWDI 60
:
DB 2 VRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLTGKGGPLPFAWDI 61
61 LSPQFGYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120
|
61 LSPQFGYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120
DB 62 LSPQFGYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 121
|
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|
DB 122 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181

QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRRHHLFL 225
|
DB 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRRHHLFL 226

RESULT 4
US-10-244-779-2
; Sequence 2, Application US/10244779
; Patent No. 6933375
; GENERAL INFORMATION:
; APPLICANT: Falcowski, Paul
; APPLICANT: Sun, Yi
; APPLICANT: Gordunov, Maxim
; APPLICANT: Wymann, Kevin
; APPLICANT: Chen, Yi-Bu
; TITLE OF INVENTION: mCFP Encoding Nucleic Acids,
; TITLE OF INVENTION: Polypeptides, Antibodies and Methods of Use Thereof
; FILE REFERENCE: Rut 00-0023US
; CURRENT APPLICATION NUMBER: US/10/244,779
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: 60/322,189
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Montastrea cavernosa
US-10-244-779-2

Query Match 56.1%; Score 681; DB 2; Length 225;
Best Local Similarity 57.2%; Pred. No. 1.7e-70;
Matches 123; Conservative 36; Mismatches 56; Indels 0; Gaps 0;

QY 6 NVIKFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLTGKGGPLPFAWDILSPQF 65
:
DB 2 SVIKSVMKIKLRMDGIYVNGHKEMITGEGGKPFEGHTILIKVKEGGPLPFAWDILTAF 61
66 QYGSKYVVKHPADIPDYKKLSPEEGFKWERVMNFEDEGVTVTQDSSLQDGCFTYKVKFI 125
|
DB 62 QYGNRVFTKYPKIDIPDYFKQSPFEGYSWERSMTFEDGVCVTITSIDIKLBGDCFFYEIRFY 121
QY 126 GVNFPSSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAK 185
|
DB 122 GVNFPSSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAK 181
QY 186 PVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGR 220
|
DB 182 GVALPEYHFVDRIEILSHDKDYNNTVEYENAVAR 216

RESULT 5
US-09-277-716-16
; Sequence 16, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238



```

; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-277-716-16
```

```

Query Match          40.0%; Score 485.5; DB 2; Length 238;
Best Local Similarity 44.5%; Pred. No. 8.5e-48;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;
```

```
QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKKLVTKGGPLPFAW 58
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTEMGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVFEFK 178
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 122 VYRVEYKGSNFPDDGPMQKTIIGIEPSFEAMVMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYVVDSKLDITSHNEDYTIYEQYE 215
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 TLMKSKGVKPEPSYHFIOHRLKKT-YVEDGGFVEQHE 218
```

```

RESULT 6
US-09-609-161B-16
; Sequence 16, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-16
```

```

Query Match          40.0%; Score 485.5; DB 2; Length 238;
Best Local Similarity 44.5%; Pred. No. 8.5e-48;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;
```

```
QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKKLVTKGGPLPFAW 58
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTEMGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVFEFK 178
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```

Db 122 VYRVEYKGSNFPDDGPMQKTIIGIEPSFEAMVMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYVVDSKLDITSHNEDYTIYEQYE 215
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 TLMKSKGVKPEPSYHFIOHRLKKT-YVEDGGFVEQHE 218
```

```

RESULT 7
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla muelleri
US-09-626-581D-65
```

```

Query Match          40.0%; Score 485.5; DB 2; Length 238;
Best Local Similarity 44.5%; Pred. No. 8.5e-48;
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;
```

```
QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKKLVTKGGPLPFAW 58
   ||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 SKQILKNTCLQEVMSYKVNLEGIVNNHVFTEMGCGKNILFGNQLVQIRVTKGAPLPFAF 61

QY 59 DILSPQFYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGC 118
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 62 DIVSPAFOYGNRTFTKYPNDISDYFIQSFPAGFMYERTLRYEDGGGLVEIRSDINLIEDKF 121

QY 119 IYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVFEFK 178
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 122 VYRVEYKGSNFPDDGPMQKTIIGIEPSFEAMVMNNGVLVGEVILVYKLSGKYYSCHMK 181

QY 179 SIYMAKKPV-QLPGYVVDSKLDITSHNEDYTIYEQYE 215
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 182 TLMKSKGVKPEPSYHFIOHRLKKT-YVEDGGFVEQHE 218
```

```

RESULT 8
US-09-415-765B-65
; Sequence 65, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A66900-1/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/415,765B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla muelleri
US-09-415-765B-65
```

```

Query Match          40.0%; Score 485.5; DB 2; Length 238;
```







RESULT 12  
US-09-977-897-14  
; Sequence 14, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 14  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-14

Query Match 39.0%; Score 473.5; DB 2; Length 225;  
Best Local Similarity 45.5%; Pred. No. 1.9e-46;  
Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

OY 5 KNV-IKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGKGPLPFAWDILSP 63  
DB 8 KNTGLKEIMSAKASVEGIVNNHVFSGMEGFGKGNVLFGNQLMQIRVTKGGPLPFAFDIVSI 67  
OY 64 QFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFIYKVK 123  
DB 68 AFQYGNRTFTKYPPDDIADYFVQSPAGFYERNLRFEDGAIVIDIRSDISLEDDKFHYKVE 127  
OY 124 FIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMA 183  
DB 128 YRNGFPSPNGPVMQKAILGMEPSFEVVMNSGVLVGEVDLVYKLESIGNYSSCHMKTFYRS 187  
OY 184 KKPV-QLPGYVVVDSKLDITSHNEDYTIYQYE 215  
DB 188 KGGVKEFPEYHFHHRLEKT-YVEEGSFVEQHE 219

RESULT 13  
US-09-977-897-15  
; Sequence 15, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluoresce  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 15  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-15

Query Match 39.0%; Score 473.5; DB 2; Length 226;  
Best Local Similarity 45.5%; Pred. No. 1.9e-46;  
Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

OY 5 KNV-IKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGKGPLPFAWDILSP 63  
DB 8 KNTGLKEIMSAKASVEGIVNNHVFSGMEGFGKGNVLFGNQLMQIRVTKGGPLPFAFDIVSI 67  
OY 64 QFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFIYKVK 123  
DB 68 AFQYGNRTFTKYPPDDIADYFVQSPAGFYERNLRFEDGAIVIDIRSDISLEDDKFHYKVE 127

OY 124 FIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMA 183  
DB 128 YRNGFPSPNGPVMQKAILGMEPSFEVVMNSGVLVGEVDLVYKLESIGNYSSCHMKTFYRS 187  
OY 184 KKPV-QLPGYVVVDSKLDITSHNEDYTIYQYE 215  
DB 188 KGGVKEFPEYHFHHRLEKT-YVEEGSFVEQHE 219

RESULT 14  
US-09-977-897-16  
; Sequence 16, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluorescer  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 16  
; LENGTH: 227  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-16

Query Match 39.0%; Score 473.5; DB 2; Length 227;  
Best Local Similarity 45.5%; Pred. No. 1.9e-46;  
Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

OY 5 KNV-IKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGKGPLPFAWDILSP 63  
DB 8 KNTGLKEIMSAKASVEGIVNNHVFSGMEGFGKGNVLFGNQLMQIRVTKGGPLPFAFDIVSI 67  
OY 64 QFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFIYKVK 123  
DB 68 AFQYGNRTFTKYPPDDIADYFVQSPAGFYERNLRFEDGAIVIDIRSDISLEDDKFHYKVE 127  
OY 124 FIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMA 183  
DB 128 YRNGFPSPNGPVMQKAILGMEPSFEVVMNSGVLVGEVDLVYKLESIGNYSSCHMKTFYRS 187  
OY 184 KKPV-QLPGYVVVDSKLDITSHNEDYTIYQYE 215  
DB 188 KGGVKEFPEYHFHHRLEKT-YVEEGSFVEQHE 219

RESULT 15  
US-09-977-897-17  
; Sequence 17, Application US/09977897  
; Patent No. 6780974  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Yih-Tai  
; APPLICANT: Cao, Longguang  
; TITLE OF INVENTION: A synthetic DNA encoding an orange seapen-derived green fluorescer  
; TITLE OF INVENTION: protein with codon preference of mammalian expression systems an  
; FILE REFERENCE: 41856-5  
; CURRENT APPLICATION NUMBER: US/09/977,897  
; CURRENT FILING DATE: 2001-10-15  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 17  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Ptilosarcus gurneyi  
US-09-977-897-17

Query Match 39.0%; Score 473.5; DB 2; Length 228;  
Best Local Similarity 45.5%; Pred. No. 2e-46;



Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

Qy	5	KNV-IKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLVTKGGPLPFAWDILSP	63
		:         :         :	
Db	8	KNTGLKEIWSAKASVEGI VNNHVFSPMEGFGNVLFGNQLMQIRVTKGGPLPFAFDIVSI	67
Qy	64	QFOYGSKYVVKHPADIPDYKCLSPPEGFKWERVMNFEEDGGVVTVTQDSSLQDGCFIYKYK	123
		: : :                     :	
Db	68	AFQYGNRFTTKYPDDIADYFVQSPFAGFFYERNLRFEDGAI VDIRSDISLEDDKFHYKVE	127
Qy	124	FIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVBFKSIYMA	183
		:                 :         :	
Db	128	YRGNGFPSPNGPVMQKAILGMEPSFEVVYMN SGVLVGBVDLVYKLESGNYYSCHMKTfYRS	187
Qy	184	KKPV-QLPGYYVDSKLDITSHNEDYTI VEQYE	215
		:         :     :     :	
Db	188	KGGVKEPPEYHFIHRL EKT-YVEEGSFVEQHE	219

Search completed: January 11, 2006, 02:05:23  
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

CM protein - protein search, using sw model

Run on: January 11, 2006, 02:00:37 ; Search time 62 Seconds  
(without alignments)  
1516.316 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQYERTGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	3	US-09-999-745-67 Sequence 67, Appl
2	1214	100.0	225	3	US-09-866-538-12 Sequence 12, Appl
3	1214	100.0	225	3	US-09-794-308-12 Sequence 12, Appl
4	1214	100.0	225	3	US-09-865-291-12 Sequence 12, Appl
5	1214	100.0	225	4	US-10-006-922-12 Sequence 12, Appl
6	1214	100.0	225	4	US-10-081-864-8 Sequence 8, Appl
7	1214	100.0	225	4	US-10-121-258-1 Sequence 1, Appl
8	1214	100.0	225	4	US-10-315-920-2 Sequence 2, Appl
9	1214	100.0	225	4	US-10-132-067-4 Sequence 4, Appl
10	1214	100.0	225	4	US-10-370-570-56 Sequence 56, Appl
11	1214	100.0	225	4	US-10-406-618-32 Sequence 32, Appl
12	1214	100.0	225	4	US-10-433-640-13 Sequence 13, Appl
13	1214	100.0	225	4	US-10-311-030-7 Sequence 7, Appl
14	1214	100.0	225	5	US-10-845-484-3 Sequence 3, Appl
15	1214	100.0	225	5	US-10-885-988-12 Sequence 12, Appl
16	1214	100.0	225	5	US-10-656-029-22 Sequence 22, Appl
17	1214	100.0	225	5	US-10-857-622-12 Sequence 12, Appl
18	1214	100.0	225	5	US-10-505-486-26 Sequence 26, Appl
19	1214	100.0	225	5	US-10-844-064A-2 Sequence 2, Appl
20	1214	100.0	225	5	US-10-931-304-1 Sequence 1, Appl
21	1214	100.0	487	4	US-10-343-977-1 Sequence 1, Appl
22	1214	100.0	487	5	US-10-343-977-1 Sequence 1, Appl
23	1214	100.0	506	4	US-10-343-977-2 Sequence 2, Appl
24	1214	100.0	506	5	US-10-343-977-2 Sequence 2, Appl
25	1214	100.0	547	4	US-10-343-977-3 Sequence 3, Appl
26	1214	100.0	547	5	US-10-343-977-3 Sequence 3, Appl
27	1210	99.7	226	4	US-10-311-030-9 Sequence 9, Appl

28	1210	99.7	240	4	US-10-152-296-2	Sequence 2, Appl
29	1210	99.7	240	4	US-10-739-656-2	Sequence 2, Appl
30	1210	99.7	545	4	US-10-214-932-52	Sequence 52, Appl
31	1210	99.7	548	4	US-10-214-932-76	Sequence 76, Appl
32	1207	99.4	225	4	US-10-121-258-20	Sequence 20, Appl
33	1207	99.4	225	4	US-10-315-920-4	Sequence 4, Appl
34	1207	99.4	225	5	US-10-931-304-20	Sequence 20, Appl
35	1196	98.5	236	4	US-10-314-936-2	Sequence 2, Appl
36	1196	98.5	236	4	US-10-314-936-4	Sequence 4, Appl
37	1196	98.5	236	6	US-11-021-014-2	Sequence 2, Appl
38	1196	98.5	236	6	US-11-021-014-4	Sequence 4, Appl
39	1195.5	98.5	226	4	US-10-724-178-12	Sequence 12, Appl
40	1191	98.1	225	4	US-10-315-920-6	Sequence 6, Appl
41	1186	97.7	225	4	US-10-442-148A-7	Sequence 7, Appl
42	1186	97.7	239	4	US-10-442-148A-8	Sequence 8, Appl
43	1160	95.6	225	4	US-10-121-258-4	Sequence 4, Appl
44	1160	95.6	225	5	US-10-931-304-4	Sequence 4, Appl
45	1153	95.0	225	4	US-10-121-258-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-09-999-745-67  
; Sequence 67, Application US/09999745  
; Patent No. US20020157120A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: REGEN1470-1  
; CURRENT APPLICATION NUMBER: US/09/999,745  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/316,920  
; PRIOR FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 67  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-999-745-67

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPEAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPEAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERYMNFEDGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWERYMNFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
DB 121 KYKFIGVNFPSDGPVQKKTGMGEASTERLYPRDGVLKGEIHKALKKDGHYLVEPKSI 180  
QY 181 YMAKKPVQLPGYTYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225  
DB 181 YMAKKPVQLPGYTYVDSKLDITSHNEDYTIIVEQYERTGRHHLFL 225

RESULT 2  
US-09-866-538-12  
; Sequence 12, Application US/09866538  
; Publication No. US20030032088A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger



```

; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-866-538-12

```

Query Match	100.0%;	Score 1214;	DB 3;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
Qy      1 MRSSKNVIKEFMRFKVRMEGTUNGHEBEIEGEGEGRPYEGHNTVKLKVTYKGGCLPFAWDI 60
        |||||||
Db      1 MRSSKNVIKEFMRFKVRMEGTUNGHEBEIEGEGEGRPYEGHNTVKLKVTYKGGCLPFAWDI 60
```

Qy 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFFEDGGVTVTTQDSSLQDGCFTY 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFFEDGGVTVTTQDSSLQDGCFTY 120

Qy 121 KVKFIGVNFPSDGPVMQKKTGWEASTERYPRDGLKEIHKALKKDGGHYLVEFKSI 180  
|||  
Db 121 KVKFIGVNFPSDGPVMQKKTGWEASTERYPRDGLKEIHKALKKDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYYYVDSKLDITSHNEDYTIIVEGYERTBGRHLLFL 225  
|||  
DB 181 YMAKKPVQLPGYYYVDSKLDTTSHNEDYTIIVEGYERTBGRHLLFL 225

### RESULT 3

```

US-09-794-308-12
; Sequence 12, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEH, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-794-308-12

```

Query Match	100.0%	Score 1214;	DB 3;	Length 225;
Best Local Similarity	100.0%	Pred. No. 6.6e-114;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	Db
1 MRSSKNVKEFMRPKVRMEGTVNGHFEIEGEGEGRPYEGHNTVKLKTKGGLPFAWDI 60	1 MRSSKNVKEFMRPKVRMEGTVNGHFEIEGEGEGRPYEGHNTVKLKTKGGLPFAWDI 60

Qy	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTTVTQDSSLQDGCFIY	120
Db	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTTVTQDSSLQDGCFIY	120

```
Qy      121 KVKFIGVNFPSDGPVMQKKTGWEASTBRLYPKDGVLKGEIHKALKLKDGGHYLVFEKSI 180
        |||||
Db      121 KVKFIGVNFPSDGPVMQKKTGWEASTBRLYPKDGVLKGEIHKALKLKDGGHYLVFEKSI 180
```

QY 181 YMAKPPVQLEPGYYYVDSKLDITSHNEDYTIIVEOYERTEGRHLLFL 225

Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHLLFL 225

## RESULT 4

US-09-865-291-12  
; Sequence 12, Application US/09865291  
; Publication No. US20030186229A1

```

; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TS'EN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-865-291-12

```

Query Match	100.0%;	Score 1214;	DB 3;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 1 MRSSKNVIEEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGGPLPFAWDI 600  
|||  
Db 1 MRSSKNVIEEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGGPLPFAWDI 600

Qy	61 LSPFOYGSKYVKHPADIPDYKLSFPEGFKMERVMNFEDGGVATYTQDSSLQDGCFTY 120 
Db	61 LSPFOYGSKYVKHPADIPDYKLSFPEGFKMERVMNFEDGGVATYTQDSSLQDGCFTY 120 

```
Oy      121 KVKFIGNFPSDGPVMQKKTGMGEASTERLYPRDGVLKEIHKALKLKDGGHYLVFEFSI   180  
        |||||  
Db      121 KVKFIGVNFPSDGPVMQKKTGMGEASTERLYPRDGVLKEIHKALKLKDDGSHYLVFEFSI   180
```

Dy 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225  
|||  
Db 181 YMAKKPVLGYYYVDSKLDTISHNEDYTIIVEQYERTEGRHHFL 225

## RESULT 5

US-10-006-922-12  
; Sequence 12, Application US/10006922  
; Publication No. US20020197676A1

```

; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Terekikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
;

```



```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma species
US-10-006-922-12
```

```
Query Match          100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60

QY      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120

QY      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||||||
DB      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||||||
DB      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 6
US-10-081-864-8
/ Sequence 8, Application US/10081864
/ Publication No. US20030022287A1
/ GENERAL INFORMATION:
/ APPLICANT: Lukyanov, Sergey
/ APPLICANT: Lukyanov, Konstantin
/ APPLICANT: Yanushevich, Yuriy
/ APPLICANT: Savitsky, Alexandr
/ APPLICANT: Fradkov, Arcady
/ TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
/ TITLE OF INVENTION: Methods for Using the Same
/ FILE REFERENCE: CLON-067
/ CURRENT APPLICATION NUMBER: US/10/081,864
/ PRIOR FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: 10/006,922
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/270,983
/ PRIOR FILING DATE: 2001-02-21
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 8
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma sp.
US-10-081-864-8
```

```
Query Match          100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60

QY      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120

QY      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||||||
DB      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||||||
```

```
DB      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 7
US-10-121-258-1
/ Sequence 1, Application US/10121258
/ Publication No. US20030059835A1
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger
/ APPLICANT: Campbell, Robert
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: UC083.1CP2CP1
/ CURRENT APPLICATION NUMBER: US/10/121,258
/ CURRENT FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 78
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(225)
/ OTHER INFORMATION: wild-type DsRed
US-10-121-258-1
```

```
Query Match          100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTYKGGLPFAWDI 60

QY      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFGYGSKVYVKGHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120

QY      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||||||
DB      121 KYKFIGVNFPSDGPVQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||||||
DB      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 8
US-10-315-920-2
/ Sequence 2, Application US/10315920
/ Publication No. US20030175809A1
/ GENERAL INFORMATION:
/ APPLICANT: Tersikh, Alexey
/ APPLICANT: Fradkov, Arcady Fedorovich
/ TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
/ TITLE OF INVENTION: FOR THEIR USE
/ FILE REFERENCE: CLON-077CIP
/ CURRENT APPLICATION NUMBER: US/10/315,920
/ CURRENT FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/211,607
/ PRIOR FILING DATE: 2000-06-14
/ PRIOR APPLICATION NUMBER: PCT/US01/19097
/ PRIOR FILING DATE: 2001-06-13
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 225
/ TYPE: PRT
```



ORGANISM: *Discosoma* sp.  
US-10-315-920-2

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Dy 1 MRSSKNVKEFMRFKVMETGVNGHFEIEGEGEGDPYEGHNTVKLKVTKGSPLPFAMD I 60

Db 1 MRSSKNVKEFMRFKVMETGVNGHFEIEGEGEGDPYEGHNTVKLKVTKGSPLPFAMD I 60

QY 61 LSPQFQYGSKVVKHPADIPDYKKLSPEEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 61 LSPQFQYGSKVVKHPADIPDYKKLSPEEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120

QY 121 KYKFIGVNPSPDGPVMQKTMTGWEASTERTLYPRDGLVKG EIHKALKLKDGGHYLVFEFKSI 180  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 121 KYKFIGVNPSPDGPVMQKTMTGWEASTERTLYPRDGLVKG EIHKALKLKDGGHYLVFEFKSI 180

QY 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHLLFL 225  
| | | | | | | | | | | | | | | | | | | | | |  
DB 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHLLFL 225

RESULT 9  
US-10-132-067-4

```

; Sequence 4, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic

```

```

; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
US-10-132-067-4

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGEGRPYEGHNTVTKLKVTGGPLPFAWDI 600
        |||||
        1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGEGRPYEGHNTVTKLKVTGGPLPFAWDI 600
        |||||
        1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGEGRPYEGHNTVTKLKVTGGPLPFAWDI 600

```

```
QY      61 LSPQFQYGSKVYVKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTTQDSSLQDGCIFY 120
      |||||
DB      61 LSPQFQYGSKVYVKHPADIPDYKCLSPFEGFKMERVMNPFEDGGVTVTTQDSSLQDGCIFY 120
```

```

Oy      121 KYKFIGVNFPSDGPVMQKKTMTGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
        |||||||
Db      121 KYKFIGVNFPSDGPVMQKKTMTGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

```

[illegible]

RESULT 10  
US-10-370-570-56  
; Sequence 56, Application US/10370570  
; Publication No. US20030219717A1

```

; GENERAL INFORMATION:
; APPLICANT: DAHL, Soren Weis et al.
; TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS
; FILE REFERENCE: 3759-0130P
; CURRENT APPLICATION NUMBER: US/10/370,570
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 71

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; US-10-370-570-56

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MRSSKNV	IK	EFMR	PK	VM	EG	TV	NG	HE	FE	IE	GE	EG	GR	PR	YE	GH	NT	YK	LK	VT	KG	GP	LP	FAM	DI	60	
Db	1	MRSSKN	VI	KE	FM	RF	KV	ME	GT	VN	GH	FE	IE	GE	EG	GR	PR	YE	GH	NT	YK	LK	VT	KG	GP	LP	FAM	DI	60

```
QY      61 LSPQFOYGSKYVVKHPADIPDYKKLSFPEGEKMERVMNFEDGGVVTVTQDSSLQDGCFIY 120
      |||||
DB      61 LSPQFOYGSKYVVKHPADIPDYKKLSFPEGEKMERVMNFEDGGVVTVTQDSSLQDGCFIY 120
```

```
Qy      121 KVKFI GVNPFSDGPMÖKKTMGWEASTERLYPRDGLKEIHKALKKDGGHYLVEFKSI    180
        |||||
Db      121 KVKFI GVNPFSDGPMÖKKTMGWEASTERLYPRDGLKEIHKALKKDGGHYLVEFKSI    180
```

[illegible]

RESULT 11  
US-10-406

```

US-10-406-618-32
; Sequence 32, Application US/10406618
; Publication No. US20030219814A1
; GENERAL INFORMATION:
; APPLICANT: Wan, David Chi-Cheong
; APPLICANT: Ip, Denis Tsz-Ming
; TITLE OF INVENTION: The Chinese University of Hong Kong
; FILE REFERENCE: No. US20030219814A1el Fluorescent Proteins
; FILE REFERENCE: 016285-34-1

```

```

;
; FURNISH EXAMINER: 02-05-97
; CURRENT APPLICATION NUMBER: US/10/406,618
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: US 60/387,968
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US 60/370,598
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentVer. 2.1

```

```

; SOFTWARE: patentin ver. 2.1
; SEQ_ID NO 32
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Spontaneously
; OTHER INFORMATION: fluorescent protein DsRed.
ITS-10-406-618-32

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	Db
1 MRSSKNVKEFMRFKVAMEGTVNGHEFEIEGEGEGRPYECHNTVYKLVTYKGGPLPFAWDI 60	1 MRSSKNVKEFMRFKVAMEGTVNGHEFEIEGEGEGRPYECHNTVYKLVTYKGGPLPFAWDI 60

```

QY      61 LSPQFOYGSKIVYKHPADIPDYKKLSFPEGEKMERVMNFEDGGVVTVTQDSSLQDGCFIY 120
        |||||||
DB      61 LSPQFOYGSKIVYKHPADIPDYKKLSFPEGEKMERVMNFEDGGVVTVTQDSSLQDGCFIY 120

```



QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 12  
US-10-433-640-13  
; Sequence 13, Application US/10433640  
; Publication No. US20040115792A1  
; GENERAL INFORMATION:  
; APPLICANT: Lichtenberg-Frate, Hella  
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF  
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS  
; FILE REFERENCE: 1487/3  
; CURRENT APPLICATION NUMBER: US/10/433,640  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/14610  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: DE 10061872.3  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discoboma sp.  
US-10-433-640-13

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 13  
US-10-311-030-7  
; Sequence 7, Application US/10311030  
; Publication No. US20040171107A1  
; GENERAL INFORMATION:  
; APPLICANT: Nelson, David  
; APPLICANT: Zamtara, Elize  
; APPLICANT: Tsielen, Roger  
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
; FILE REFERENCE: 15916-032US1  
; CURRENT APPLICATION NUMBER: US/10/311,030  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: PCT/US01/04625  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US 60/184,732  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7

; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-311-030-7

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 14  
US-10-845-484-3  
; Sequence 3, Application US/10845484  
; Publication No. US20040248180A1  
; GENERAL INFORMATION:  
; APPLICANT: Bulina, Maria  
; APPLICANT: Chudakov, Dmitry  
; APPLICANT: Lukanov, Konstantin  
; TITLE OF INVENTION: Mutant Chromaphores/Fluorophores and  
; TITLE OF INVENTION: Methods for Making and Using the Same  
; FILE REFERENCE: CLON 092  
; CURRENT APPLICATION NUMBER: US/10/845,484  
; PRIOR FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: 60/343128  
; PRIOR FILING DATE: 2001-12-26  
; PRIOR APPLICATION NUMBER: PCT/US02/41418  
; PRIOR FILING DATE: 2002-12-23  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-10-845-484-3

Query Match 100.0%; Score 1214; DB 5; Length 225;  
Best Local Similarity 100.0%; Pred. No. 6.6e-114;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225



```

RESULT 15
US-10-885-988-12
; Sequence 12, Application US/10885988
; Publication No. US20040259165A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEH, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/10/885,988
; CURRENT FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: US/09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Dicobosoma sp.
US-10-885-988-12

```

Query Match		100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity		100.0%;	Pred. No. 6.6e-114;		
Matches 225;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAMD	60		
Db	1	MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAMD	60		
QY	61	LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFFEDGGVTVTQDSSLQDGCFTY	120		
Db	61	LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFFEDGGVTVTQDSSLQDGCFTY	120		
QY	121	KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVPEKSI	180		
Db	121	KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVPEKSI	180		
QY	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225		
Db	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225		

Search completed: January 11, 2006, 02:06:37  
Job time : 63 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:03:07 ; Search time 8 Seconds  
(without alignments)  
238.884 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTGRHLLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 61141

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	6	US-10-209-208-1 Sequence 1, Appli
2	1207	99.4	225	6	US-10-209-208-20 Sequence 20, Appli
3	1160	95.6	225	6	US-10-209-208-4 Sequence 4, Appli
4	1153	95.0	225	6	US-10-209-208-24 Sequence 24, Appli
5	1121	92.3	226	6	US-10-209-208-6 Sequence 6, Appli
6	1021	84.1	225	6	US-10-209-208-8 Sequence 8, Appli
7	1012	83.4	225	6	US-10-209-208-79 Sequence 79, Appli
8	485.5	40.0	238	7	US-11-179-411-16 Sequence 16, Appli
9	485.5	40.0	238	7	US-11-175-766-16 Sequence 16, Appli
10	473.5	39.0	238	7	US-11-179-411-32 Sequence 32, Appli
11	473.5	39.0	238	7	US-11-175-766-32 Sequence 32, Appli
12	422.5	34.8	233	7	US-11-179-411-27 Sequence 27, Appli
13	422.5	34.8	233	7	US-11-175-766-27 Sequence 27, Appli
14	410.5	33.8	233	7	US-11-179-411-33 Sequence 33, Appli
15	410.5	33.8	233	7	US-11-175-766-33 Sequence 33, Appli
16	210.5	17.3	238	6	US-10-209-208-10 Sequence 10, Appli
17	210.5	17.3	239	7	US-11-089-551A-40 Sequence 40, Appli
18	209.5	17.3	239	6	US-10-209-208-12 Sequence 12, Appli
19	205.5	16.9	239	6	US-10-209-208-15 Sequence 15, Appli
20	204.5	16.8	512	7	US-11-032-236-6 Sequence 6, Appli
21	204.5	16.8	1219	7	US-11-032-236-4 Sequence 4, Appli
22	200.5	16.5	239	6	US-10-209-208-13 Sequence 13, Appli
23	192.5	15.9	239	6	US-10-209-208-11 Sequence 11, Appli
24	192.5	15.9	239	6	US-10-209-208-14 Sequence 14, Appli
25	183.5	15.1	437	7	US-11-069-642-92 Sequence 92, Appli

26	182.5	15.0	438	7	US-11-069-642-47	Sequence 47, Appli
27	182.5	15.0	438	7	US-11-069-642-49	Sequence 49, Appli
28	182.5	15.0	438	7	US-11-069-642-51	Sequence 51, Appli
29	182.5	15.0	438	7	US-11-069-642-53	Sequence 53, Appli
30	182.5	15.0	438	7	US-11-069-642-55	Sequence 55, Appli
31	182.5	15.0	438	7	US-11-069-642-57	Sequence 57, Appli
32	182.5	15.0	438	7	US-11-069-642-59	Sequence 59, Appli
33	182.5	15.0	438	7	US-11-069-642-61	Sequence 61, Appli
34	182.5	15.0	438	7	US-11-069-642-63	Sequence 63, Appli
35	103	8.5	559	6	US-10-793-626-1376	Sequence 1376, Ap
36	80	6.6	424	6	US-10-467-657-90	Sequence 90, Appli
37	80	6.6	424	6	US-10-467-657-90	Sequence 90, Appli
38	80	6.6	690	7	US-11-052-554A-232	Sequence 232, App
39	76.5	6.3	325	6	US-10-467-657-5568	Sequence 5568, Ap
40	76.5	6.3	325	6	US-10-467-657-5568	Sequence 5568, Ap
41	76.5	6.3	574	7	US-11-053-100-40	Sequence 40, Appli
42	76.5	6.3	724	7	US-11-053-100-41	Sequence 41, Appli
43	76.5	6.3	874	7	US-11-053-100-42	Sequence 42, Appli
44	76.5	6.3	1174	7	US-11-053-100-43	Sequence 43, Appli
45	76.5	6.3	1663	6	US-10-055-877-148	Sequence 148, App

ALIGNMENTS

RESULT 1  
US-10-209-208-1  
; Sequence 1, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(225)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-1

Query Match 100.0%; Score 1214; DB 6; Length 225;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTNGHBEIEGEGRPYEGHNTVKLKVTGSGPLPFAWDI	60
DB	1	MRSSKNVKEFMRFKVRMEGTNGHBEIEGEGRPYEGHNTVKLKVTGSGPLPFAWDI	60
QY	61	ISPQFQYGSKVYKHPADIPDYKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGCFTY	120
DB	61	ISPQFQYGSKVYKHPADIPDYKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGCFTY	120
QY	121	KYKFIGVNFPSDGPVWQKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI	180
DB	121	KYKFIGVNFPSDGPVWQKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI	180
QY	181	YNAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTGRHLLFL	225



|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 2

US-10-209-208-20  
; Sequence 20, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dated with 1125R  
US-10-209-208-20

Query Match 99.4%; Score 1207; DB 6; Length 225;  
Best Local Similarity 99.6%; Pred. No. 1.1e-109;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLKVTGKGPLPFAWDI 60  
|||  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
|||  
Db 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
Db 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3

US-10-209-208-4  
; Sequence 4, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26

; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dated polypeptide variant "T1"  
US-10-209-208-4

Query Match 95.6%; Score 1160; DB 6; Length 225;  
Best Local Similarity 96.0%; Pred. No. 3.7e-105;  
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLKVTGKGPLPFAWDI 60  
|||  
Db 1 MASSEDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQYAKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
|||  
Db 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
Db 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 4

US-10-209-208-24  
; Sequence 24, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dated polypeptide variant "T1" with 1125R mutation  
US-10-209-208-24

Query Match 95.0%; Score 1153; DB 6; Length 225;  
Best Local Similarity 95.6%; Pred. No. 1.8e-104;  
Matches 215; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVYKLKVTGKGPLPFAWDI 60  
|||  
Db 1 MASSEDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQYAKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
|||  
Db 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180



```
Db      121 KVKFRGVNFPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKSI 180
QY      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 5
US-10-209-208-6
; Sequence 6, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "dimer2"
US-10-209-208-6
```

```
Query Match      92.3%; Score 1121; DB 6; Length 226;
Best Local Similarity 92.8%; Pred. No. 2.2e-101;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY      3 SSKNVKEFMRFKVRMEGTVNGHFEIEEGEGRPYEGHNTVKLKVTKGGLPFAWDILS 62
Db      4 SSEDVKEFMRFKVRMEGSVNGHFEIEEGEGRPYEGTQAKLKVTKGGLPFAWDILS 63

QY      63 PQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGVVTVTQDSSLQDGCFTYKV 122
Db      64 PQFQYGSKAYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGVVTVTQDSSLQDGTLLYKV 123

QY      123 KEIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKSIYM 182
Db      124 KFRGTNFPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKTIYM 183

QY      183 AKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      184 AKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERSEGRHHLFL 226
```

```
RESULT 6
US-10-209-208-8
; Sequence 8, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
```

```
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mRFP1"
US-10-209-208-8
```

```
Query Match      84.1%; Score 1021; DB 6; Length 225;
Best Local Similarity 86.9%; Pred. No. 1e-91;
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db      1 MASSEDVKEFMRFKVRMEGSVNGHFEIEEGEGRPYEGTQAKLKVTKGGLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFQYGSKAYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGVVTVTQDSSLQDGEFTY 120

QY      121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKSI 180
Db      121 KVKLRGTNFPDGPVMQKTMGWEASTERMYPEDGALKGEIKRLKDKGHYDAEVKTT 180

QY      181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRH 221
Db      181 YMAKKPVQLPGAAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 221
```

```
RESULT 7
US-10-209-208-79
; Sequence 79, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mRFP1.1"
US-10-209-208-79
```

```
Query Match      83.4%; Score 1012; DB 6; Length 225;
Best Local Similarity 86.0%; Pred. No. 7.4e-91;
Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db      1 MASSEDVKEFMRFKVRMEGSVNGHFEIEEGEGRPYEGTQAKLKVTKGGLPFAWDI 60

QY      61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGVVTVTQDSSLQDGCFTY 120
```



Db 61 LSPQFWYGSKAYVKHPADIPDYLKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGEFII 120  
QY 121 KVKFIGNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKLRGTFNFPDGPVMQKTMGWEASERMYPEDGALKGEIKRLLKDKGGHYDAEVKTI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 181 YMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 221

RESULT 8  
US-11-179-411-16  
; Sequence 16, Application US/11179411  
; Publication No. US20050266491A1  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TH  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/179,411  
; CURRENT FILING DATE: 2005-07-12  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 08/757,046  
; PRIOR FILING DATE: 1996-11-25  
; PRIOR APPLICATION NUMBER: 08/597,274  
; PRIOR FILING DATE: 1996-02-06  
; PRIOR APPLICATION NUMBER: 08/908,909  
; PRIOR FILING DATE: 1997-08-08  
; PRIOR APPLICATION NUMBER: 08/990,103  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Renilla mulleri  
US-11-179-411-16

Query Match 40.0%; Score 485.5; DB 7; Length 238;  
Best Local Similarity 44.5%; Pred. No. 5.8e-40;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAW 58  
Db 2 SKQILKNTCLQEWMSYKVNLEGIVNNHVFTMEGCGKNILFGNLVQIRVTKGAPLPFAF 61  
QY 59 DILSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGEF 118  
Db 62 DIVSPAQYGNRTFTKYPNDISDYFIQSPAGFMYERTLRYEDGGLVEIRSDINLIEDKF 121  
QY 119 IYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFK 178  
Db 122 VYRVEYKGSNFPDDGPVMQKTIIGIEPSFEAMYMNGVLVGEVILVYKLN SGKYYSCMHK 181  
QY 179 SIYMAKKPV-QLPGYYVDSKLDITSHNEDYTIIVEQYE 215  
Db 182 TLMKSKGVVKEFPSPYHFIOHRLKKT-YVEDGGFVEQHE 218

RESULT 9  
US-11-175-766-16  
; Sequence 16, Application US/11175766  
; Publication No. US20050272111A1

; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TH  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/175,766  
; CURRENT FILING DATE: 2005-07-06  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 08/757,046  
; PRIOR FILING DATE: 1996-11-25  
; PRIOR APPLICATION NUMBER: 08/597,274  
; PRIOR FILING DATE: 1996-02-06  
; PRIOR APPLICATION NUMBER: 08/908,909  
; PRIOR FILING DATE: 1997-08-08  
; PRIOR APPLICATION NUMBER: 08/990,103  
; PRIOR FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Renilla mulleri  
US-11-175-766-16

Query Match 40.0%; Score 485.5; DB 7; Length 238;  
Best Local Similarity 44.5%; Pred. No. 5.8e-40;  
Matches 97; Conservative 43; Mismatches 71; Indels 7; Gaps 3;  
QY 4 SKNVIK-----EFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAW 58  
Db 2 SKQILKNTCLQEWMSYKVNLEGIVNNHVFTMEGCGKNILFGNLVQIRVTKGAPLPFAF 61  
QY 59 DILSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGEF 118  
Db 62 DIVSPAQYGNRTFTKYPNDISDYFIQSPAGFMYERTLRYEDGGLVEIRSDINLIEDKF 121  
QY 119 IYKVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFK 178  
Db 122 VYRVEYKGSNFPDDGPVMQKTIIGIEPSFEAMYMNGVLVGEVILVYKLN SGKYYSCMHK 181  
QY 179 SIYMAKKPV-QLPGYYVDSKLDITSHNEDYTIIVEQYE 215  
Db 182 TLMKSKGVVKEFPSPYHFIOHRLKKT-YVEDGGFVEQHE 218

RESULT 10  
US-11-179-411-32  
; Sequence 32, Application US/11179411  
; Publication No. US20050266491A1  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: Szczepaniak, William  
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING TH  
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-128  
; CURRENT APPLICATION NUMBER: US/11/179,411  
; CURRENT FILING DATE: 2005-07-12  
; PRIOR APPLICATION NUMBER: US/09/808,898  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/189,691  
; PRIOR FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 09/277,716  
; PRIOR FILING DATE: 1999-03-26



```

; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Ptilosarcus gurneyi
; OS-11-179-411-32

```

Query Match	39.0%;	Score 473.5;	DB 7;	Length 238;
Best Local Similarity	45.5%;	Pred. No. 8.3e-39;		
Matches 97;	Conservative 40;	Mismatches 73;	Indels 3;	Gaps 3;

Qy	5	KNV- I KEFMR FKVRMEGTAVNGHEEIEEGEGECR PYREGHNTVKLKVTKGGPLPFAMDILSP	63
Dd	7	KNTGLKEIMSAKASVEGIVNNHVSMEGFGKNVLFGNQLMQIRVTKGGPLPFAFDIVI	66
Qy	64	QEFGSKVYVKHPADIPDYKKLSFPPEGFKMERVMNFEDGGVTVTQDSLODGCFIYKVK	123
Dd	67	AFOYGNRFTTKYPDDIADYFVGSPAGFFYERNLRFEDEGAIVDIRSDISLEDDKHYYKE	126
Qy	124	PICGNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHAKLKLXDGHYLVEFSIYMA	183
Dd	127	YRGNGFPSNGPVMQKAILGMPEPSFEVVYMNSSGVLVGEVDLVYKLESNGNYSSCHMKTFFYS	186
Qy	184	KKPV-QLPGYYYYVDSKLDITSHNEDYTIIVEQYE	215
Dd	187	KGVKKEFFEYHFIIHRLEKT-YVEEGSFVEOHE	218

RESULT 11  
US-11-175

Sequence 32, Application US/11175766  
Publication No. US2005027211A1

```

; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szczepaniak, William
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING 1
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/11/175,766
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: US/09/808,898
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32

```

TYPE: PRT  
ORGANISM: *Ptilobarbus gurneyi*  
US-11-175-766-32

```

Query Match          39.0%; Score 473.5; DB 7; Length 238;
Best Local Similarity 45.5%; Pred. No. 8.3e-39;
Matches 97; Conservative 40; Mismatches 73; Indels 3; Gaps 3;

OY 5 KNV-IKEFMRFKVRMEGTVNGHFEIEGEGRPRYEGHNTVKLKVTKGGLPFAWDILSP 63
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 KNTGLKEIMSAKASVEGI VNNHVFSGMEFGKGNVLFGNQLMQIRVTKGGPLPFAFDIVSI 66

OY 64 QFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGCVTVTQDSSLQDGCFFIYKV 123
   ||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 67 AFQYGNRTFTYKPPDDIADYFVQSPFAGFFYERNLRFEDGAI VDIRSDISLEDKFKHYKVE 126

OY 124 FIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEKSIYMA 183
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 127 YRGNGPSNGPVMQKAILGMEPSFEVVMNSGVLVGEVDLVYKLESNGNYTSCHMKTFYRS 186

OY 184 KKPV-QLPGYVYVDSKLDITSHNEDYTIYQYE 215
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 187 KGVGKEFPEYHFIRHLEKT-YVEGGSFVEQHE 218

RESULT 12
US-11-179-411-27
; Sequence 27, Application US/11179411
; Publication No. US20050266491A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; TITLE OF INVENTION: SZCZEPANIAK, WILLIAM
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/11/179,411
; CURRENT FILING DATE: 2005-07-12
; PRIOR APPLICATION NUMBER: US/09/808,898
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Renilla reniformis
US-11-179-411-27

Query Match          34.8%; Score 422.5; DB 7; Length 233;
Best Local Similarity 40.5%; Pred. No. 6.9e-34;
Matches 90; Conservative 38; Mismatches 93; Indels 1; Gaps 1;

OY 1 MSSSKNVIKEFMRFKVRMEGTVNGHFEIEGEGRPRYEGHNTVKLKVTKGGLPFAWDI 60
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 1 MDLAKLGLKEVMPTKINLEGLVGDHAFSGMEVGEGNILEGTQEVKISVTKGAPLPFAFDI 60

OY 61 LSPQFQYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGCVTVTQDSSLQDGCFFIY 120
   |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 61 VSAFSGYNRAYTGYPEEISDYPLQSPFEGFTYERNIRYQDGTATVKS DISLEDGKFIV 120

OY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEKSI 180
   ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 121 NVDFEAKADLRMGPMQDVGMPDSYESMYTNTVTSVIGECIIAFLKQTGKHFTYHMRIV 180

```







```

; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Renilla Reniformis muten
US-11-175-766-33

```

ORGANISM: Ren11a Reniformis muteln  
US-11-175-766-33

Query Match	33.8%	Score 410.5;	DB 7;	Length 233;
Best Local Similarity	40.1%;	Pred. No. 9.9e-33;		
Matches 89;	Conservative 37;	Mismatches 95;	Indels 1;	Gaps 1;

```

Qy      1 MRSSKAVIKEFMRFKVRMEGTVNGHFEIEEGEGGRPYEGHNTVTKLKYTKGSPLPFAMDI 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MDLAKDLGIKEVMPYTKINLEGLVGDBHAFSMEGVGEGNILEGTQEVKISVTKGAPLPFAFDI 60

Qy      61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEEGFKMERVMNFEDGVVTYTQDSSLQDGCFTY 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 VSAFSGYNRAYTGYPEEISDYFLQSFPEEGFTYERNIRYQDGGTAIVSDISLEDGKFIV 120

Qy      121 KYKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKG EIHKALKLKDGGHYLVEFKSI 180
      | | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121 NVDFKADDLRDMGPVMQDIVGMQPSYSEMYTNVTNVTSVIGECIIAFKLQTKDFTYHMRIV 180

Qy      181 YMAKKPVQ-LPGYVVVDSKLDITSHNEDYTIVEQYERTEGRH 221
      | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181 YKSKKPVETMPLVYHFIOHDLVKTNVDTASGVVVQHETAIAAH 222

```

Search completed: January 11, 2006, 02:06:56  
Job time : 8 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n\_model

Run on: January 12, 2006, 09:05:37 ; Search time 3691 Seconds  
(without alignments)  
3465.125 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US10006922/runat\_10012006\_162408\_18921/app\_query.fasta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922 @CGN\_1\_1\_7415 @runat\_10012006\_162408\_18921 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	6	AR183915 Sequence
2	1214	100.0	678	6	AX172854 Sequence
3	1214	100.0	678	6	AX207715 Sequence

4	1214	100.0	678	6	AX233581	AX233581 Sequence
5	1214	100.0	678	6	AX370404	AX370404 Sequence
6	1214	100.0	678	6	AX824725	AX824725 Sequence
7	1214	100.0	859	2	AF168419	AF168419 Discosoma
8	1214	100.0	859	6	AR636082	AR636082 Sequence
9	1214	100.0	859	6	AR670157	AR670157 Sequence
10	1214	100.0	859	6	AX463698	AX463698 Sequence
11	1210	99.7	681	6	AX233584	AX233584 Sequence
12	1210	99.7	713	6	AX233627	AX233627 Sequence
13	1210	99.7	723	6	AR527331	AR527331 Sequence
14	1210	99.7	2721	6	CQ882115	CQ882115 Sequence
15	1210	99.7	2772	6	CQ882117	CQ882117 Sequence
16	1210	99.7	4692	6	AX463702	AX463702 Sequence
17	1210	99.7	6893	6	AX823860	AX823860 Sequence
18	1210	99.7	8811	11	AY569779	AY569779 Cloning v
19	1210	99.7	9320	6	AX663075	AX663075 Sequence
20	1210	99.7	12404	11	AY569780	AY569780 Cloning v
21	1207	99.4	678	6	AX370406	AX370406 Sequence
22	1196	98.5	666	6	AX348043	AX348043 Sequence
23	1196	98.5	666	6	AX353910	AX353910 Sequence
24	1196	98.5	711	2	AY679107	AY679107 Discosoma
25	1193	98.3	678	2	AF545838	AF545828 Discosoma
26	1191	98.1	678	6	AX370408	AX370408 Sequence
27	1191	98.1	678	6	AX824732	AX824732 Sequence
28	1191	98.1	921	2	AY679106	AY679106 Discosoma
29	1186	97.7	675	6	AX824731	AX824731 Sequence
30	1186	97.7	1050	6	AX666133	AX666133 Sequence
31	1186	97.7	4488	6	CQ849509	CQ849509 Sequence
32	1186	97.7	4546	11	AY818375	AY818375 Cloning v
33	1186	97.7	4555	11	DQ005468	DQ005468 Cloning v
34	1186	97.7	4570	11	AY818373	AY818373 Cloning v
35	1186	97.7	4829	11	AY640628	AY640628 SiRNA vec
36	1186	97.7	4906	11	AY640625	AY640625 SiRNA vec
37	1186	97.7	5311	11	AJ851284	AJ851284 Cloning v
38	1186	97.7	6423	11	AY613997	AY613997 Cloning v
39	1186	97.7	6990	11	AY640634	AY640634 SiRNA vec
40	1186	97.7	7058	11	AY640633	AY640633 SiRNA vec
41	1186	97.7	7147	11	AY640630	AY640630 SiRNA vec
42	1186	97.7	7616	6	CQ849511	CQ849511 Sequence
43	1186	97.7	9468	6	CS018252	CS018252 Sequence
44	1186	97.7	10140	6	CS018253	CS018253 Sequence
45	1186	97.7	10522	6	CS018254	CS018254 Sequence

ALIGNMENTS

RESULT 1	AR183915	AR183915	678 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR183915	Sequence 6 from patent US 6342379.				
DEFINITION	AR183915					
ACCESSION	AR183915					
VERSION	AR183915.1	GI:20227884				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 678)					
AUTHORS	Tsien,R.Y. and Gonzalez,J.E. III.					
TITLE	Detection of transmembrane potentials by optical methods					
JOURNAL	Patent: US 6342379-A 6 29-JAN-2002;					
FEATURES	Location/Qualifiers					
source	1..678					
	/organism="unknown"					
	/mol_type="unassigned DNA"					

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114  
Score: 1214.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 6  
Length: 678  
Matches: 225  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0



US-10-006-922A-12 (1-225) x ARI83915 (1-678)

QY	1	MetArGSeSerSerLySAsnValIllelysgLuPheMeCArgPheLySValArgMeGlugly	20
Db	1	ATGAGGTCTTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGGTTCCATGGAAGGA	60
QY	21	ThrValAsnGlyHlsgLuPheGluIllegluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCGACGAGTTTGAAATGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC	120
QY	41	HisAsnThrValLySLeuLySValThrLysgLyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAATACCGTTAAAGCTTAAGGTAACCAAGGGGGACCTTGCCATTGGCTTGGAATAT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyshIsProAlaAspIlePro	80
Db	181	TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATTAATAAGCTGTCAATTCTCGAAGGATTTAAATGGGAAAGGGTCATGAACCTTGA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTyr	120
Db	301	GACGGTGGCGTCTGACTGTAACCAAGATTCCAGTTGCAGATGGCTGTTTCATCTAC	360
QY	121	LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
Db	361	AAGGTCAAGTTCAATTGGCGTGAACCTTCCGATGACCTGTTATGCAAAAGAGACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysgLyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGGTTTGATCTCGTATGGCGGTGGAAGAGAG	480
QY	161	IleHisLySAIalaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle	180
Db	481	ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTTAAGTTCAAAAAGTATT	540
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
Db	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGACGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCATCTGTTCCCTT 675	
RESULT 2			
LOCUS	AX172854	678 bp	DNA linear PAT 03-JUL-2001
DEFINITION	Sequence 6 from Patent WO0142211.		
ACCESSION	AX172854		
VERSION	AX172854.1	GI:14597903	
KEYWORDS	. Discosoma sp.		
SOURCE	Discosoma sp.		
ORGANISM	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.		
REFERENCE	1	Tsien,R.Y. and Gonzalez,J.E.	
AUTHORS	Detection of transmembrane potentials by optical methods		
TITLE	Patent: WO 0142211-A 6 14-JUN-2001;		
JOURNAL	The Regents of the University of California (US)		
FEATURES	location/Qualifiers		
SOURCE	1..678		
	/organism="Discosoma sp."		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:86600"		
ORIGIN			

Alignment Scores:

Pred. No.:	1.4e-114	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AX172854 (1-678)

QY	1	MetArGSeSerSerLySAsnValIllelysgLuPheMeCArgPheLySValArgMeGlugly	20
Db	1	ATGAGGTCTTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGGTTCCATGGAAGGA	60
QY	21	ThrValAsnGlyHlsgLuPheGluIllegluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCGACGAGTTTGAAATGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC	120
QY	41	HisAsnThrValLySLeuLySValThrLysgLyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAATACCGTTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyshIsProAlaAspIlePro	80
Db	181	TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTGCCGACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATTAATAAGCTGTCAATTCTCGAAGGATTTAAATGGGAAAGGGTCATGAACCTTGA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTyr	120
Db	301	GACGGTGGCGTCTGACTGTAACCAAGATTCCAGTTGCAGATGGCTGTTTCATCTAC	360
QY	121	LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
Db	361	AAGGTCAAGTTCAATTGGCGTGAACCTTCCGATGACCTGTTATGCAAAAGAGACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysgLyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGGTTTGATCTCGTATGGCGGTGGAAGAGAG	480
QY	161	IleHisLySAIalaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle	180
Db	481	ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTTAAGTTCAAAAAGTATT	540
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
Db	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGACGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCATCTGTTCCCTT 675	
RESULT 3			
LOCUS	AX207715	678 bp	DNA linear PAT 31-AUG-2001
DEFINITION	Sequence 13 from Patent WO0157242.		
ACCESSION	AX207715		
VERSION	AX207715.1	GI:15422399	
KEYWORDS	. Discosoma sp.		
SOURCE	Discosoma sp.		
ORGANISM	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.		
REFERENCE	1	Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.	
AUTHORS	Methods of protein destabilization and uses thereof		
TITLE			



```

JOURNAL      Patent: WO 0157242-A 13 09-AUG-2001,
              Aurora Biosciences Corporation (US)
FEATURES
  source      location/Qualifiers
              1. . 678
              /organism="Discosoma sp."
              /mol_type="unassigned DNA"
              /db_xref="taxon:86600"

```

QY	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

```

KEYWORDS
SOURCE
ORGANISM
1. Discosoma sp.
Discosoma sp.
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE
AUTHORS
1 Nelson, D., Zamatra, E. and Tsien, R.
TITLE
Modified fluorescent proteins
JOURNAL
Patent: WO 0162919-A 5 30-AUG-2001;
Aurora Biosciences Corporation (US)
FEATURES
source
Location/Qualifiers
1. .678
/organism="Discosoma sp."
/mol_type="unassigned DNA"
/db_xref="taxon:86600"
/note="red"

```

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGAGGCTCTCCAAGAATGTTATCAAGAGGTTCAAGAGTTTAAGTTTCGATGGAAGGA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCATGGGCACGACGAGTTGAATAAGAAGCGAAGAGAGAGGGGCCATACGAAGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCTTGGCAATTGCTGGGATATT	180
QY	61	LeuSerProGluPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	TTGTCAACCACAATTTCAGTATGAGACCAAGTATATGCAAGCACCTTCGCCACATACCA	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATAAAAGCTGTCAATTTCCGTAAGATTTAATGGAAAGGGTCATGAACTTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGTCGTTACTGTAAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTCAAGTTCATTGGCGTGAACTTTCCTCCGATGACCTGTTATGCAAAAGAAGACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCCTCGTATGCGGTGTGAAGAAGAGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATTAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAAGACCGAGGACGC	660
QY	221	HisHisLeuPheLeu	225



Db	661	CACCATCTGTTCCCTT	675	
RESULT 5				
LOCUS	AX370404	678 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0196373.			
ACCESSION	AX370404			
VERSION	AX370404.1	GI:18857490		
KEYWORDS				
SOURCE	synthetic construct			
ORGANISM	synthetic construct			
REFERENCE	1	other sequences; artificial sequences.		
AUTHORS	Fradekov, A.F. and Tersikh, A.			
TITLE	Fluorescent timer proteins and methods for their use			
JOURNAL	Patent: WO 0196373-A 1 20-DEC-2001;			
FEATURES	Clontech Laboratories Inc. (US)			
source	Location/Qualifiers			
	1. 678			
	/organism="synthetic construct"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:32630"			
	/note="variant of sequence from Discosoma sp."			
ORIGIN				
Alignment Scores:				
Pred. No.:	1.4e-114	Length:	678	
Score:	1214.00	Matches:	225	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	6	Gaps:	0	
US-10-006-922A-12 (1-225) x AX370404 (1-678)				
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20	
DB	1	ATGCGCTCTCCAGAACGTCAATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGGC	60	
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40	
DB	61	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGC	120	
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60	
DB	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTCGCTTCCGCTGGACATC	180	
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80	
DB	181	CTGTCCCCCAGTCCAGTACGAGCTCCAGAGGTGATGAGACACCCCGGACATCCCC	240	
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100	
DB	241	GACTACAGAAGCTGTCTTCCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG	300	
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120	
DB	301	GACGGCGGCGTGTGACCGTGACCCAGGACTCTCTGACGAGACGGCTGTCTTCACTAC	360	
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140	
DB	361	AAGTGAAGTTCATCGCGGTGAACCTTCCCTCCGAGGCCCCCGTGATGACAGAGAACCC	420	
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160	
DB	421	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGGACGCGCTGTGAAGGGGAG	480	
QY	161	IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180	
DB	481	ATCCACAGGCGCTGAAGCTGAAGAGACGGCGGCCACTACCTGTGTGAGTTCAAGTTCATC	540	
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200	

Db	541	TACATGGCCAGAGACCCCGTGCAGCTGCGCGGCTACTACTACGTGACGTCCAGCTGGAC	600
Qy	201	ILEThrSerHisAsnGluAspTyrThrILEValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACCTCCCAACAGAGACTACACCATCTGTGAGCAGTACGAGCCGACCGAGGCCCC	660
Qy	221	HISHisLeuPheLeu	225
Db	661	CACCACCTGTTCTGTG	675
RESULT 6			
AX824725		678 bp	DNA
LOCUS			linear
DEFINITION	Sequence 7 from Patent WO02068459.		PAT 11-DEC-2003
AX824725			
AX824725	GI:39750591		
VERSION			
KEYWORDS			
SOURCE	Discosoma sp.		
ORGANISM	Discosoma sp.		
REFERENCE			
AUTHORS	1		
TITLE	Non aggregating fluorescent proteins and methods for using the same		
JOURNAL	Patent: WO 02068459-A 7 06-SEP-2002;		
FEATURES	Location/Qualifiers		
source	1..678		
	/organism="Discosoma sp."		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:86600"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.4e-114	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AX824725 (1-678)			
Qy	1	MetArgSerSerIysAsnValILEValGluPheMetArgPheIysValArgMetGluGly	20
Db	1	ATGCGCTCTCCCAAGACGTATCAAGAGTTTCATCGCTTCAAGGTGCGCATGGAGGGC	60
Qy	21	ThrValAsnGlyHisGluPheGluILEGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACAGATTGAGATGAGGCGGAGGGCGAGGGCCGCCCTACGAGGGC	120
Qy	41	HISAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheIATrPApIle	60
Db	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCTGCTCGCTGGGACATC	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCAGATCCCC	240
Qy	81	AspTyrIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAGAACTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAG	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGGCGTGTGACCGTGACCCAGGACTCTCCCTGACGAGGCGGCTGCTCATCTAC	360
Qy	121	LysValIysPheILEGlyValAsnPheProSerAspGlyProValMetGlnIysLysThr	140
Db	361	AAAGTGAAGTTTCATCGCGGTGAATTCCTCCGACGCGCCCGTGATGCAAGAAAGACC	420
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu	160



Db	421	ATGGCTGGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGCGGAG	480
Qy	161	IlEhIsLySaLaLeuLySLeuLySaPglYgLyHIsTyrrLeuValGluphElySserIle	180
Db	481	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTCGTGGAGTTCAAGTCCATC	540
Qy	181	TyrMetAlaLySLeYsProValGlnLeuProGlyTyrrTyrrValAspSerLySLeuAsp	200
Db	541	TACATGGCCCAAGAAGCCCGTGCAGCTGCCGCTACTACTACGTGACTCCAAGCTGAC	600
Qy	201	IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrrGluArgThrGluGlyArg	220
Db	601	ATCACCTTCCACAACGAGACTACACCATCTGTGAGCAGTACGAGCCGACCGAGGCGCC	660
Qy	221	HisHisLeuPheLeu	225
Db	661	CACCACCTGTCTCTG	675
RESULT 7			
LOCUS	AF168419	859 bp	mRNA linear INV 27-JUL-2001
DEFINITION	Discosoma sp. fluorescent protein FP583	mRNA, complete cds.	
ACCESSION	AF168419		
VERSION	AF168419.2	GI:7105733	
KEYWORDS	Discosoma sp.		
SOURCE	Discosoma sp.		
ORGANISM	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomacidae; Discosoma.		
REFERENCE	1 (bases 1 to 859)		
AUTHORS	Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.		
TITLE	Fluorescent proteins from nonbioluminescent Anthozoa species		
JOURNAL	Nat. Biotechnol. 17 (10), 969-973 (1999)		
PUBMED	10504696		
REFERENCE	2 (bases 1 to 859)		
AUTHORS	Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.		
TITLE	Direct Submision		
JOURNAL	Submitted (14-JUL-1999) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia		
REFERENCE	3 (bases 1 to 859)		
AUTHORS	Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.		
TITLE	Direct Submision		
JOURNAL	Submitted (25-FEB-2000) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia		
REMARK	Sequence update by submitter		
COMMENT	On Feb 25, 2000 this sequence version replaced gi:6090866.		
FEATURES	Location/Qualifiers		
source	1..859		
	/organism="Discosoma sp."		
	/mol_type="mRNA"		
	/db_xref="taxon:86600"		
	54..731		
	/note="GFP-like protein; orange-red"		
	/codon_start=1		
	/product="Fluorescent protein FP583"		
	/protein_id="AAF03369.1"		
	/db_xref="GI:6090867"		
	/translation="MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTV KLKVTKGGLPFPADILSPFOYGSKVYVKHPADIPDYKLSFPEGFKERVNPNFEDG GVVTVQDSSLQDGCFLYKVKFIGNFSDGPVMQKKTMGWEASTERLYPRDGLKGE IHKALKDKDGHYLVPEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTE GRHHLFL"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.87e-114	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB:	2	Gaps:	0
US-10-006-922A-12 (1-225) x AF168419 (1-859)			
Qy	1	MetArgSerSerLySAsnValIleIySgluPhemEtaArgPheLySValArgMetGluGly	20
Db	54	ATGAGGCTTCCAAGAATGTTATCAAGAGTTTCATGAGGTTTAAGTTCCGATGGAAGCA	113
Qy	21	ThrValAsnGlyHIsGluPheGluIleGluGlyGluGlyArgProTyrrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTGAAATGAAAGCGGAAGGAGAGGGGAGCCATACGAAGCC	173
Qy	41	HisAsnThrValLySLeuLySValThrLySGLYGLYProLeuProPheAlaTrpAspIle	60
Db	174	CACAAATCCGTAAAGCTTAAGGTAAACCAAGGGGACCCTTGCCATTGCTGGATATT	233
Qy	61	LeuSerProGlnPheGlnTyrrGlySerLySValTyrrValLySHisProAlaAspIlePro	80
Db	234	TTGTCAACCACAATTTCAGTATGGAAGCATATATGTCAAGCACCCCTGCGACATACCA	293
Qy	81	AspTyrrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAAGCTGTCAATTCTTCGAGCAATTTAAATGGGAAAGGTCATGAACTTGAA	353
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrr	120
Db	354	GACGGTGGCTGCTTACTGTAAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC	413
Qy	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThr	140
Db	414	AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGACCTGTATGCAAAAGAAAGACA	473
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValIleuLySGLYGLU	160
Db	474	ATGGCTGGAGACCAAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAAGAGAG	533
Qy	161	IlEhIsLySaLaLeuLySLeuLySaPglYgLyHIsTyrrLeuValGluphElySserIle	180
Db	534	ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCATTACCTAGTTGAATCAAAAAGTATT	593
Qy	181	TyrMetAlaLySLeYsProValGlnLeuProGlyTyrrTyrrValAspSerLySLeuAsp	200
Db	594	TACATGGCAAGAGCCTGTGACGCTACCGAGTACTATATGTTGACTTCAAACCTGGAT	653
Qy	201	IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGACGC	713
Qy	221	HisHisLeuPheLeu	225
Db	714	CACCATCTGTCTCTT	728
RESULT 8			
LOCUS	AR636082	859 bp	DNA linear PAT 14-FEB-2005
DEFINITION	Sequence 11 from patent US 6852849.		
ACCESSION	AR636082		
VERSION	AR636082.1	GI:59795931	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 859)		
AUTHORS	Tsien,R.Y. and Campbell,R.E.		
TITLE	Non-oligomerizing tandem fluorescent proteins		
JOURNAL	Patent: US 6852849-A 11 08-FEB-2005;		
FEATURES	The Regents of the University of California; Oakland, CA		
source	1..859		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			



Alignment Scores:

Pred. No.:	1.87e-114	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AR636082 (1-859)

QY	1	MetArGSeSerLySAsnValIlElYsGlUPhemeArGPhelYsValArgMetGlUGly	20
DB	54	ATGAGGTCCTCCAGAATGTTATCAAGAGTTCAATGAGGTTTAAGGTTCCGCAATGGAAGCA	113
QY	21	ThrValAsnGlYhIsGlUPhEgLuIeGlUGlYgLUglYArgProTYrGluGly	40
DB	114	ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGGGAGGCCCATACGAAGGC	173
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTYrGlySerLySValTYrValLySHisProAlaAspIlePro	80
DB	234	TTGTCAACCACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	293
QY	81	AspTYrLyLySLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPheGlu	100
DB	294	GACTATAAAAGCTGTCTATTCCGTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTYr	120
DB	354	GACGGTGGCGTCGTTACTGTACCAAGATTCAGATTTCAGAGATGGCTGTTTCATCTAC	413
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
DB	414	AAAGTCAAGTTCATTGGCGTGAACTTTCCTCCGATGACCTGTTATGCCAAAAGAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLySGlyGlu	160
DB	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGAAGCGCTGTGAAAGGAGAG	533
QY	161	IlEhIsLySAlaLeuLySLeuLySAspGlyGlyhIsTYrLeuValGluPheLySserIle	180
DB	534	ATTCAATAAGGCTCTGAAGCTGAAGAAGGCTGTCATTACCTAGTTGAATCAAAAAGTATT	593
QY	181	TYrMetAlaLySLeSProValGlnLeuProGlyTYrTYrTYrValAspSerLySLeuAsp	200
DB	594	TACATGGCAAGAAGCCTGTGCAAGTACCAAGGTACTACTATGTGAACTCCAAACTGGAT	653
QY	201	IlEthrSerHisAsnGluAspTYrThrIleValGlnGlnTYrGluArgThrGluGlyArg	220
DB	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGAGCGC	713
QY	221	HisHisLeuPheLeu 225	
DB	714	CACCATCTGTTCCCTT 728	

RESULT 9  
 AR670157  
 LOCUS AR670157 859 bp DNA linear PAT 13-JUN-2005  
 DEFINITION Sequence 11 from patent US 6900304.  
 ACCESSION AR670157  
 VERSION AR670157.1 GI:67608896  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Unkown.  
 Unclassified.

REFERENCE 1 (bases 1 to 859)  
 AUTHORS Tsien,R.Y., Ting,A.Y. and Zhang,J.  
 TITLE Emission ratiometric indicators of phosphorylation  
 JOURNAL Patent: US 6900304-A 11 31-MAY-2005;  
 The Regents of the University of California, Oakland, CA

FEATURES  
 Source Location/Qualifiers  
 1..859  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	1.87e-114	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AR670157 (1-859)

QY	1	MetArGSeSerLySAsnValIlElYsGlUPhemeArGPhelYsValArgMetGlUGly	20
DB	54	ATGAGGTCCTCCAGAATGTTATCAAGAGTTCAATGAGGTTTAAGGTTCCGCAATGGAAGCA	113
QY	21	ThrValAsnGlYhIsGlUPhEgLuIeGlUGlYgLUglYArgProTYrGluGly	40
DB	114	ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGGGAGGCCCATACGAAGGC	173
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTYrGlySerLySValTYrValLySHisProAlaAspIlePro	80
DB	234	TTGTCAACCACAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	293
QY	81	AspTYrLyLySLeuSerPheProGluGlyPheLySTrPGluArgValMetAsnPheGlu	100
DB	294	GACTATAAAAGCTGTCTATTTCCTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTYr	120
DB	354	GACGGTGGCGTCGTTACTGTAAACCAAGATTCAGATTTCAGAGATGGCTGTTTCATCTAC	413
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
DB	414	AAAGTCAAGTTCATTGGGCTGAACCTTCCGATGAGCCTGTATGCAAAAAGAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLySGlyGlu	160
DB	474	ATGGGCTGGGAAGCCAGCACTGAAGCTTTGTATCCTCGTGAATGGCGTGTGAAAGGAGAG	533
QY	161	IlEhIsLySAlaLeuLySLeuLySAspGlyGlyhIsTYrLeuValGluPheLySserIle	180
DB	534	ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTTAAGTTGAAATCAAAAAGTATT	593
QY	181	TYrMetAlaLySLeSProValGlnLeuProGlyTYrTYrTYrValAspSerLySLeuAsp	200
DB	594	TACATGGCAAGAAGCCTGTGCAAGTACCAAGGTACTACTATGTGACTCCAAACTGGAT	653
QY	201	IlEthrSerHisAsnGluAspTYrThrIleValGlnGlnTYrGluArgThrGluGlyArg	220
DB	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGAGCGC	713
QY	221	HisHisLeuPheLeu 225	
DB	714	CACCATCTGTTCCCTT 728	

RESULT 10  
 AX463698  
 LOCUS AX463698 859 bp DNA linear PAT 15-JUL-2002  
 DEFINITION Sequence 12 from Patent WO0248338.  
 ACCESSION AX463698  
 VERSION AX463698.1 GI:21886457  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Discosoma sp.  
 Discosoma sp.



REFERENCE  
AUTHORS  
TITLE  
1 Lichtenberg-Frat, H.  
Yeast strain for testing the geno- and cytotoxicity of complex environmental contamination

JOURNAL  
Patent: WO 0248338-A 12 20-JUN-2002;  
Lichtenberg-Frat, Hella (DE)  
Location/Qualifiers

FEATURES  
Source  
1. .859  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
54. .731  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD42147.1"  
/db\_xref="GI:21886458"  
/translation="MRSSKNVIKEFMRFKVRMEGTUNGHEFEIEGEGRPYEGHNTV  
KLKVTKGGLPFAWDILSPQFGSKYVKHPADIPDYKLSFPEGKMERVMNPFEDG  
GVVTVTQDSSLQDGCIFYKVKFIGNFSPSDGPMQKTMGEASTERLYPRDGLKGE  
IHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTE  
GRHHLFL"

CDS  
1.87e-114  
1214.00  
100.00%  
100.00%  
100.00%  
6  
859  
225  
0  
0  
0  
0  
0

ORIGIN

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:  
US-10-006-922A-12 (1-225) x AX463698 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20  
DB 54 ATGAGGCTTCCAGAAATGTTATCAAGAGATTCAAGGTTTAAGGTTCCATGGAAGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCAAGATTGAAATAGAAAGGCCAAGAGAGGGGAGGCCATA CGAAGGC 173  
QY 41 HisAsnThrValLySLeuLySValThrLySglYProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAAGCTTAAAGTAAACCAAGGGGGACCTTTGCCATTGCTTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 294 GACTATAAAAAGCTGTCATTTCTCGAAGGATTAAATGGGAAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGTGGCGTCGTTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTCACTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
DB 414 AAGGTCAAGTTCAATGGCGTGAACTTTCTTCCGATGGACCTGTTATGCAAAAGAGACA 473  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProAlrAspGlyValLeuLySglYglu 160  
DB 474 ATGGGCTGGGAAGCCAGCACTGACGCTTGTATCTCGATGATGGCGTGTGAAGAGAG 533  
QY 161 IleHisLySAIalaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180  
DB 534 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTTAAGTCAAAAGTATT 593  
QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

DB 594 TACATGGCAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAACTGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATAACAAGCCACAACGAAAGACTATACAATCGTTGAGCAGTATGAAGAACGAGGAGCGC 713  
QY 221 HisHisLeuPheLeu 225  
DB 714 CACCATCTGTTCTT 728

RESULT 11  
AX233584  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
11  
AX233584  
Sequence 8 from Patent WO0162919.  
AX233584  
AX233584.1 GI:15593307  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
1 Nelson, D., Zamaira, E. and Tsien, R.  
Modified fluorescent proteins  
Patent: WO 0162919-A 8 30-AUG-2001;  
Aurora Biosciences Corporation (US)  
Location/Qualifiers

FEATURES  
Source  
1. .681  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"  
1. .681  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC69734.1"  
/db\_xref="GI:15593308"  
/translation="MRSSKNVIKEFMRFKVRMEGTUNGHEFEIEGEGRPYEGHNT  
VRLKVTKGGLPFAWDILSPQFGSKYVKHPADIPDYKLSFPEGKMERVMNPFEDG  
GVVTVTQDSSLQDGCIFYKVKFIGNFSPSDGPMQKTMGEASTERLYPRDGLKGE  
EIHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERT  
EGRHHLFL"

CDS

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:  
US-10-006-922A-12 (1-225) x AX233584 (1-681)

QY 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20  
DB 4 GTGAGGAGCAGCAAGACGTGATCAAGAGTTCAAGGTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCACGAGTTCAGATCGAGGCGGAGGGCGAGGCCCTACGAGGGC 123  
QY 41 HisAsnThrValLySLeuLySValThrLySglYglYProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAGCTTAAAGTGACCAAGGGCGGCCCTGCTGCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 184 CTGAGCCCCCAGTTCCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCGCATCCCC 243  
QY 81 AspTyrLySlySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAGAAGCTGAGCTTCCCCGAGGCTTCAAGTGGAGAGGCTGATGAACCTTCGAG 303



QY

101

ASPG1YGLYVALValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

DB

304

GACGGCGCGGTGTGTGACCGTGACCCAGACAGACAGCCCTGTGAGACGGCTGCTTCATCTAC

363

QY

121

LYSValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr

140

DB

364

AAAGTGAAGTTCATCGCGCTGAACCTTCCCGACGACGCGCCCGTGATGCAGAAAGAACCC

423

QY

141

METGlyTRPG1uAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

DB

424

ATGGGCTGGGAGGCGCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG

483

QY

161

ILEH1SLYSAlaLeuLysLeuLysAspGlyGlyHIS1TYrLeuValGluPheLysSerIle

180

DB

484

ATCCACAAAGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGGAGTTCAGTCCATC

543

QY

181

TYRMETAlaLysLysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAsp

200

DB

544

TACATGGCCCAAGAACCCCGTGCAGCTGCCGCTACTACTAGTGAAGTCCAAAGCTGAC

603

QY

201

ILEThrSerHISAsnGluAspTYrThrIleValGluGlnTYrGluArgThrGluGlyArg

220

DB

604

ATCACCAAGCCCAACGAGACTACCATCGTGAGCAGTACGAGAGAACCGAGGGCGAGG

663

QY

221

HIS1SLeuPheLeu 225

225

DB

664

CACCACCTGTTCCTG 678

678

RESULT 12

AX233627

LOCUS

AX233627

DEFINITION

Sequence 51 from Patent WO0162919.

ACCESSION

AX233627

VERSION

AX233627.1

KEYWORDS

GI:15593330

SOURCE

synthetic construct

ORGANISM

synthetic construct

REFERENCE

1

AUTHORS

Nelson,D., Zamaira,E. and Tsien,R.

TITLE

Modified fluorescent proteins

JOURNAL

Patent: WO 0162919-A 51 30-AUG-2001;

FEATURES

Aurora Biosciences Corporation (US)

source

location/Qualifiers

1..713

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Mutant Anchozoan red fluorescent protein"

ORIGIN

Alignment Scores:

Pred. No.:

3.81e-114

Score:

1210.00

Length:

713

Matches:

224

Percent Similarity:

100.00%

Conservative:

1

Best Local Similarity:

99.56%

Mismatches:

0

Query Match:

99.67%

Indels:

0

DB:

6

Gaps:

0

US-10-006-922A-12 (1-225) x AX233627 (1-713)

QY

1

MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly

20

DB

22

GTGAGGACGACGCAAGACGTGATCAAGAGTTCATGAGGTTCAAGGTGCGCATGAGGGC

81

QY

21

ThrValAsnGlyHISGluPheGluIleGluGlyGluGlyArgProTYrGluGly

40

DB

82

ACCGTGAACGGCCACGAGTTTCAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGGC

141

QY

41

HISAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle

60

DB

142

CACAAACACCGTGAAGCTTAAAGTGACCAAGGGCGGCCCTGCTTGGCTGGGACATC

201

QY

61

LeuSerProGlnPheGlnTYrGlySerLysValTYrValLysHISProAlaAspIlePro

80

DB

202

CTGAGCCCCCAGATTCCAGTACGGCAAGAGGTGTACGTGAAGCACCCCCGACATCCCC

261

QY

81

ASPTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu

100

DB

262

GACTACAAAGAGCTGAGCTTCCCGAGGGCTTCAAGTGGAGAGAGGTGATGAACCTTGAG

321

QY

101

ASPG1YGLYVALValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

DB

322

GACGGCGCGGTGTGACCGGTGACCCAGACAGACAGCCTGCAGAGCGGCTGCTCATCTAC

381

QY

121

LYSValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr

140

DB

382

AAAGTGAAGTTCATCGCGCTGAACCTTCCCGACGACGCGCCCGTGATGCAGAAAGAAC

441

QY

141

METGlyTRPG1uAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

DB

442

ATGGGCTGGAGGCTTCACCGAGCGCCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG

501

QY

161

ILEH1SLYSAlaLeuLysLeuLysAspGlyGlyHIS1TYrLeuValGluPheLysSerIle

180

DB

502

ATCCACAAAGCCCTGAAGCTGAAGACGCGCGCCACTAAGTGTGAGTTCAGTTCATC

561

QY

181

TYRMETAlaLysLysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAsp

200

DB

562

TACATGGCCCAAGAACCCCGTGCAGCTGCCGCTACTACTAGTGAAGTCCAAAGCTGAC

621

QY

201

ILEThrSerHISAsnGluAspTYrThrIleValGluGlnTYrGluArgThrGluGlyArg

220

DB

622

ATCACCAAGCCCAACGAGACTACCATCGTGAGCAGTACGAGAGAACCGAGGGCGAGG

681

QY

221

HIS1SLeuPheLeu 225

225

DB

682

CACCACCTGTTCCTG 696

696

RESULT 13

AR527331

LOCUS

AR527331

DEFINITION

Sequence 1 from patent US 6723537.

ACCESSION

AR527331

VERSION

AR527331.1

KEYWORDS

GI:53914309

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1

AUTHORS

Unclassified.

TITLE

1 (bases 1 to 723)

JOURNAL

Directed evolution of protein in mammalian cells

Patent: US 6723537-A 1 20-APR-2004;

FEATURES

Rigel Pharmaceuticals, Incorporated and Becton, Dickinson and

source

Company; South San Francisco, CA

location/Qualifiers

1..723

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:

3.88e-114

Score:

1210.00

Length:

723

Matches:

224

Percent Similarity:

100.00%

Conservative:

1

Best Local Similarity:

99.56%

Mismatches:

0

Query Match:

99.67%

Indels:

0

DB:

6

Gaps:

0

US-10-006-922A-12 (1-225) x AR527331 (1-723)

QY

1

MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly

20

DB

4

GTCCGCTCCTCCAAAGACGTATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC

63



QY	21	ThrValAsnGlyH1sGluPheGluIleGluGlyGlyGlyArgProTyrGluGly	40
DB	64	ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCCGCCCTACGAGGGC	123
QY	41	HisAsnThrValIleuLeuLysValThrIysGlyGlyProLeuProPheAlaTrpAspIle	60
DB	124	CACAACACCGTGAAGCTGAAGGTGACCAAGGGGGCCCCCTGCCCTTGGGACATC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro	80
DB	184	CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCCGACATCCCC	243
QY	81	AspTyrIleLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
DB	244	GACTACAAAGAGTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG	303
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	304	GACGCGCGCGTGTGACCGTGACCCAGACTCTCCCTGCAGGACGGCTGCTTCATCTAC	363
QY	121	LysValIleLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
DB	364	AAAGTGAAGTTCAATCGCGGTGAACCTCCCTCCGACGGCCCCGTAATGCAAGAAAGACC	423
QY	141	MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu	160
DB	424	ATGGGCTGGAGGCGCTTCCACCGAGCGCTGTACCCCCCGCAGCGCGCTGTGAAGGGCGAG	483
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
DB	484	ATCCACAAGGCGCTGAAGCTGAAGAGACGGCGGCCACTAAGTGTGAGTTCAAGATATC	543
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	544	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTAAGTGAAGTCCAGCTGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	604	ATCACTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGACCGAGGGCGGC	663
QY	221	HisHisLeuPheLeu	225
DB	664	CACCACTGTTCCTG	678

RESULT 14  
 LOCUS CQ882115 2721 bp DNA linear PAT 11-OCT-2004  
 DEFINITION Sequence 1 from Patent WO2004083445.  
 ACCESSION CQ882115  
 VERSION CQ882115.1 GI:54034825  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

1  
 Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,  
 Julien,B., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.  
 Nucleic acid controlling the expression of a useful polypeptide in  
 the posterior silk glands of a lepidoptera and application thereof  
 Patent: WO 2004083445-A 1 30-SEP-2004;  
 Centre National De La Recherche Scientifique-CNRS (FR); UNIVERSITE  
 CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche  
 Agronomique (INRA) (FR)  
 FEATURES  
 SOURCE  
 1..2721  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Description de la s quence artificielle : s quence  
 de fusion"

Alignment Scores:

Pred. No. :	1.95e-113	Length:	2721
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x CQ882115 (1-2721)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
DB	2044	GTGCGCTCTCCAAAGACGTCAATCAAGAGTTGATGCGCTTCAAGTGCATGAGGGC	2103
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyArgProTyrGluGly	40
DB	2104	ACCGTGAACGGCCACGAGTTCGAGTCCAGGGCGAGGGCGAGGGCCCCCTACGAGGC	2163
QY	41	HisAsnThrValIleuLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
DB	2164	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTGGGACATC	2223
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro	80
DB	2224	CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGACATCCCC	2283
QY	81	AspTyrIleLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
DB	2284	GACTACAAAGACTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG	2343
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	2344	GACGCGCGCGTGTGACCGTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAAAGACC	2403
QY	121	LysValIleLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
DB	2404	AAAGTGAAGTTCAATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAAAGACC	2463
QY	141	MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu	160
DB	2464	ATGGGCTGGAGGCGCTTCCACCGAGCGCTGTACCCCCGAGCGCGCTGTGAAGGGCGAG	2523
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
DB	2524	ATCCACAAGGCGCTGAAGCTGAAGAGCGGGCCACTAAGTGTGAGTTCAAGTCCATC	2583
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	2584	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTAAGTGAAGTCCAGCTGAC	2643
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	2644	ATCACCTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCGGC	2703
QY	221	HisHisLeuPheLeu	225
DB	2704	CACCACTGTTCCTG	2718

RESULT 15  
 LOCUS CQ882117 2772 bp DNA linear PAT 11-OCT-2004  
 DEFINITION Sequence 3 from Patent WO2004083445.  
 ACCESSION CQ882117  
 VERSION CQ882117.1 GI:54034827  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

1  
 Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B.,  
 Julien,B., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C.  
 Nucleic acid controlling the expression of a useful polypeptide in  
 the posterior silk glands of a lepidoptera and application thereof  
 Patent: WO 2004083445-A 3 30-SEP-2004;



Centre National De La Recherche Scientifique-CNRS (FR) ; UNIVERSITE  
CLAUDE BERNARD - LYON 1 (FR) ; Institut National de la Recherche  
Agronomique (INRA) (FR)  
Location/Qualifiers  
1..2772  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description de la s quence artificielle : s quence  
de fusion"

ORIGIN

Alignment Scores:			
Pred. No.:	2e-113	Length:	2772
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x CQ882117 (1-2772)

QY	1	MeArGSeSerLySAAnValIleLySGluPhMeMetArGPhelysValArgMetGluGly	20
DB	2095	GTGGCTCTCCCAAGACGTGATCAAGAGTTCATGCGCTCAAGGTGCGCATGAGGGC	2154
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	2155	ACCGTGAACGGCCACGAGTTCGATCGAGGGCGAGGGCGGCCCTTACGAGGGC	2214
QY	41	HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	2215	CACAAACCCGTGAAGCTGAAGTGACCAAGGGCGGCCCTTGCCTGGGACATC	2274
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValIlyHisProAlaAspIlePro	80
DB	2275	CTGTCCCCCAGTTCAGTACGGCTCCAAGGTGTAGTGAACCCCGCGACATCCCC	2334
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPhGlu	100
DB	2335	GACTACAAAGAGTGTCTCTCCCGAGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	2394
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr	120
DB	2395	GACGGCGGCTGTGACCGTGACCCAGACTCTCTCCGCGAGACGGCTGCTCATCTAC	2454
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
DB	2455	AAAGTGAAGTTCATCGGCGGTGAATTCCCTCCGACGGCCCCGTAATGAGAGAAGACC	2514
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlu	160
DB	2515	ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCCCGGACGGCGTGTGAAGGGCGAG	2574
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180
DB	2575	ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGTGAGTTCAAGTTCATC	2634
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	2635	TACATGGCCAAAGACCCGTGCGCTGCCCGCTACTACTACGTGGACTCCAAGCTGAGC	2694
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	2695	ATCACTCCCAACAAGAGGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGGC	2754
QY	221	HisIleLeuPheLeu	225
DB	2755	CACCACTGTTCCTG	2769



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 05:49:00 ; Search time 479 Seconds  
(without alignments)  
3130.597 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTYVEQYERTGRHLLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USFTO\_spool\_p/US10006922/runat\_10012006\_162408\_18913/app\_query.fasta\_1.391  
-DB=N Geneseg -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN\_1\_1\_1096 @runat\_10012006\_162408\_18913 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseg\_21: \*  
1: genesegn1980s: \*  
2: genesegn1990s: \*  
3: genesegn2000s: \*  
4: genesegn2001as: \*  
5: genesegn2001bs: \*  
6: genesegn2002as: \*  
7: genesegn2002bs: \*  
8: genesegn2003as: \*  
9: genesegn2003bs: \*  
10: genesegn2003cs: \*  
11: genesegn2003ds: \*  
12: genesegn2004as: \*  
13: genesegn2004bs: \*  
14: genesegn2005s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1214	100.0	678	4	AAH47654	Aah47654 Discosoma
2	1214	100.0	678	4	AAD13053	Aad13053 Discosoma
3	1214	100.0	678	4	AAD11142	Aad11142 Discosoma
4	1214	100.0	678	6	ABA95905	Aba95905 Yeast opt

5	1214	100.0	678	6	ABA95922	Abas95922 Yeast opt
6	1214	100.0	678	6	ABA95921	Abas95921 Yeast opt
7	1214	100.0	678	6	ABA95920	Abas95920 Yeast opt
8	1214	100.0	678	6	AAD46278	Aad46278 Discosoma
9	1214	100.0	678	6	AAD28207	Aad28207 Discosoma
10	1214	100.0	678	10	ADC24127	Adc24127 Discosoma
11	1214	100.0	678	10	ADF70404	Adf70404 Discosoma
12	1214	100.0	678	11	ADL46204	Adl46204 Discosoma
13	1214	100.0	678	11	ADN33978	Adn33978 Wild-type
14	1214	100.0	678	12	ADJ36420	Adj36420 Discosoma
15	1214	100.0	678	12	ADM97768	Adm97768 D bp red
16	1214	100.0	695	3	AAA48743	Aaa48743 Humanised
17	1214	100.0	859	6	AAL47952	Aal47952 Discosoma
18	1214	100.0	859	7	ADY51715	Ady51715 Discosoma
19	1214	100.0	859	8	AAD53432	Aad53432 Discosoma
20	1214	100.0	859	10	AAD61969	Aad61969 Discosoma
21	1214	100.0	859	14	ADX26533	Adx26533 Discosoma
22	1214	100.0	3311	9	ACA62895	Aca62895 Plasmid D
23	1210	99.7	681	10	ADC24134	Adc24134 Discosoma
24	1210	99.7	681	11	ADL46205	Adl46205 Human cod
25	1210	99.7	681	11	ADL46225	Adl46225 Human cod
26	1210	99.7	723	8	ABZ22476	Abz222476 Mammalian
27	1210	99.7	1638	10	ADL18131	Adl18131 RFP:PS(NI
28	1210	99.7	1647	10	ADL18155	Adl18155 RFP:PS(HI
29	1210	99.7	2721	13	ADS75466	Ads75466 Fibrohexa
30	1210	99.7	2772	13	ADS75468	Ads75468 Fibrohexa
31	1210	99.7	4692	6	AAL47954	Aal47954 Modified
32	1210	99.7	4692	10	ACC44640	Acc44640 Vector pd
33	1210	99.7	5436	4	AAD10003	Aad10003 Plasmid p
34	1210	99.7	6893	10	ADE24111	Ade24111 Proviral
35	1210	99.7	9910	4	AAD09979	Aad09979 pBit(dhsp
36	1210	99.7	9320	6	ABS56664	Abs56664 Plasmid p
37	1210	99.7	9658	12	ADJ04104	Adj04104 Vector pT
38	1210	99.7	9678	12	ADJ04103	Adj04103 Vector pT
39	1210	99.7	10263	12	ADJ04102	Adj04102 Vector pT
40	1209	99.6	681	6	ABA95906	Abas95906 Yeast opt
41	1209	99.6	681	6	ABA95907	Abas95907 Yeast opt
42	1207	99.4	678	6	AAD28208	Aad28208 Discosoma
43	1205	99.3	681	4	AAH47656	Aah47656 Anthozoan
44	1203	99.1	7508	13	ADR30814	Adr30814 Zebrafish
45	1197	98.6	6985	10	ACA55359	Aca55359 Transform

ALIGNMENTS

RESULT 1	
AAH47654	
ID	AAH47654 standard; cDNA; 678 BP.
AC	AAH47654;
XX	
DT	30-NOV-2001 (first entry)
XX	
DE	Discosoma sp. red drFP583 protein coding sequence.
XX	
KW	Fluorescent protein; Anthozoan; fluorescence; marker; FRET; drFP583; ss.
XX	
OS	Discosoma sp.
XX	
PN	WO200162919-A1.
XX	
PD	30-AUG-2001.
XX	
PF	13-FEB-2001; 2001WO-US004625.
XX	
PR	23-FEB-2000; 2000US-0184732P.
XX	
PA	(AURO-) AURORA BIOSCIENCES CORP.
XX	
PI	Nelson D, Zamaira E, Tsien R;
XX	
DR	WPI; 2001-557704/62.
XX	



PT Proteins for fluorescence Resonance Energy Transfer (FRET) comprise  
PT functional red fluorescent proteins, and the encoding nucleic acids, with  
PT key mutations for improving the proteins function.  
XX  
PS

PS Disclosure; Page 83; 90pp; English.

CC The invention provides a nucleic acid encoding functional red fluorescent  
CC protein (II) that differs from the sequence of an Anthozoan red  
CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents the coding sequence of a Discosoma sp. drFP583 protein, an  
CC anthozoan fluorescent protein  
XX  
SQ

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-006-922A-12 (1-225) x AAH47654 (1-678)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAGATGTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGAGAGGGAGGCCATACGAAGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGGCATTTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 181 TTGTACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAAGCTGTCAATTCTCGAAGGATTTAATGGGAAAGGTCATGAACTTTGA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCGTACTGTAAACCAAGATTCCAGTTTGACGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGCGCGTGTGAAAGAGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGACGGGTGTCATTACTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAAAGAACCTGTGCAAGCTAACAGGGTACTACTATGTGTGACTCCAAACTGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAAACAAGCCACAAACGACGAAGACTATACAATCGTTGACAGATATGAAAGAACCGAGGACGC 660

QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTCTT 675

RESULT 2  
AAD13053  
ID AAD13053 standard; cDNA; 678 BP.  
XX  
AC AAD13053;

DT 16-OCT-2001 (first entry)

DE Discosoma sp. "red" anthozoa fluorescent protein, drFP583 cDNA.

KW protein destabilisation; linker moiety; reporter moiety; disease model;  
KW linear multimerised domain; -NH-ubiquitin protein endoprotease;  
KW transgenic animal; transgenic plant; disease resistance;  
KW anthozoa fluorescent protein; natural fluorescent protein; ss.

XX Discosoma sp.

XX WO200157242-A2.

PD 09-AUG-2001.

PF 02-FEB-2001; 2001WO-US003791.

PR 04-FEB-2000; 2000US-00498098.

PA (AURO-) AUFORA BIOSCIENCES CORP.

PI Stack JH, Whitney M, Cubitt AB, Pollok BA;

XX WPI; 2001-488890/53.

PT Destabilizing proteins in living cells, by coupling a target protein to  
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin  
PT protein endoproteases, comprising two copies of the domain.

PS Disclosure; Page 110; 171pp; English.

XX The present invention relates to a method for destabilising a target  
CC protein in a cell. The method comprises a linker moiety which operatively  
CC couples a target protein (a reporter moiety) to a linear multimerised  
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein  
CC endoproteases. The method is useful for detecting an activity such as  
CC protease, protein kinase or phosphoprotein phosphatase activity and is  
CC also useful for identifying modulators of these activities. The method is  
CC also useful for developing novel assays for a wide range of post-  
CC translational activities, such as proteolysis, phosphorylation,  
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,  
CC disulfide bond formation and ADP-ribosylation within cells. The  
CC recombinant DNA molecule of the invention is useful for creating  
CC transgenic animals useful as disease models and transgenic plants with  
CC improved disease resistance or other favourable traits. The present  
CC sequence is Discosoma sp. "red" anthozoa fluorescent protein, drFP483  
CC cDNA which is a natural fluorescent protein used as a reporter moiety in  
CC the exemplification of the invention  
XX  
SQ

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0



US-10-006-922A-12 (1-225) x AAD13053 (1-678)

OY 1 MetArgSerSerIysAsnValIleIySGluPheMetArgPheIysValArgMetGluGly 20  
DB 1 ATGAGGCTTCCAGAAATGTTATCAAGGAGTTCAAGGTTTAAGGTTCCGATGGAAGGA 60  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGAGAGGGAGGCCATACGAAGGC 120  
OY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGACCCTTGCCATTGCTTGGAATATT 180  
OY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
OY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCAATTCTCTGAAGATTTAATGGAAAGGGTCATGAACCTTGAA 300  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCCAGGATCCAGTTTGACAGGATGGCTTTTCATCTAC 360  
OY 121 IysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
DB 361 AAGGTCAAGTTCATTGGCGTGAACTTCTCCGATGGAACCTGTTATGCAAAAGAAAGACA 420  
OY 141 MetGlyTyrGluIleAsnThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCCTCGTGATGGCGTTGAAAGAGAG 480  
OY 161 IleHisIysValIleLeuIysLeuIysAspGlyIysHisTyrLeuValGluPheIysSerIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGTCAATTAACCTAGTTGAATTCAAAAGTATT 540  
OY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
DB 541 TACATGGCAGAAGAGCTGTGCAGCTACCAAGGGTACTACTATGTGACTCCAAACTGGAT 600  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACACAAGACTATACAATCGTTGACGATGATAAAGAACCGAGGACGC 660  
OY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTCTT 675  
RESULT 3  
AAD11142  
ID AAD11142 standard; DNA; 678 BP.  
XX AAD11142;  
AC AAD11142;  
XX  
DT 24-SEP-2001 (first entry)  
XX  
DE Discosoma species Anthozoa fluorescent protein, drFP583 encoding DNA.  
XX  
KM Transmembrane potential; biological membrane; fluorescent ion; detection;  
KM test chemical screening; Anthozoa fluorescent protein; FP;  
KM transgenic organism; drFP583 protein; ds.  
XX  
OS Discosoma sp.  
XX  
PN WO200142211-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 12-DEC-2000; 2000WO-US033739.  
XX

PR 13-DEC-1999; 99US-00459956.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Tsien RY, Gonzalez JB;  
PI  
XX WPI, 2001-457276/49.  
DR

XX  
PT Determining electrical potential across a membrane in biological systems,  
PT comprises introducing two reagents, exposing the membrane to light and  
PT measuring the energy transfer.

XX  
PS Disclosure; Page 150; 154pp; English.

XX  
CC The patent discloses optical methods and compositions for determining  
CC transmembrane potential across biological membranes in living cells. The  
CC method of determining the electrical potential across a membrane in a  
CC biological system comprises introduction of two reagents, a first reagent  
CC comprising a hydrophobic fluorescent ion capable of redistributing from a  
CC first face of the membrane to a second face of the membrane in response  
CC to membrane potential change and a second reagent that label the first  
CC face or the second face of the membrane, which comprises a chromophore  
CC capable of undergoing energy transfer by either donating or accepting  
CC excited state energy to the fluorescent ion. The membrane is then exposed  
CC to excitation light and the energy transfer between the reagents is  
CC measured and related to the membrane potential. The method is useful for  
CC detecting changes in membrane potential in subcellular organelle  
CC membranes in biological systems. The method is used for screening of test  
CC chemicals for activity to modulate the activity of target ion channel.  
CC The invention also provides a transgenic organism comprising a first  
CC reagent that comprises a charged hydrophobic fluorescent molecule and a  
CC second reagent comprising a bioluminescent or naturally fluorescent  
CC protein. The present sequence is Discosoma species (red) DNA encoding an  
CC Anthozoa fluorescent protein (FP), drFP583  
XX

SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-006-922A-12 (1-225) x AAD11142 (1-678)

OY 1 MetArgSerSerIysAsnValIleIySGluPheMetArgPheIysValArgMetGluGly 20  
DB 1 ATGAGGCTTCCAGAAATGTTATCAAGGAGTTCAAGGTTTAAGGTTCCGATGGAAGGA 60  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGAGAGGGAGGCCATACGAAGGC 120  
OY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGACCCTTGCCATTGCTTGGAATATT 180  
OY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
OY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCAATTCTCTGAAGATTTAATGGAAAGGGTCATGAACCTTGAA 300  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCCAGGATCCAGTTTGCAAGATGGCTTTTCATCTAC 360  
OY 121 IysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140



Db 361 AAGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAAGACA 420  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAATGGCGTGTGAAGAAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAGTATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATTAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCAGAGGACGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTCTCCTT 675  
RESULT 4  
ID ABA95905 standard; cDNA; 678 BP.  
XX ABA95905;  
AC ABA95905;  
XX 29-MAY-2002 (first entry)  
DT 29-MAY-2002 (first entry)  
DE Yeast optimised RFP encoding cDNA SEQ ID NO 1.  
XX  
KW Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
XX Anthozoa.  
OS  
XX  
FH Key location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "yeast optimised Red Fluorescent Protein"  
FT  
XX  
PN DE20001395-U1.  
XX  
PD 15-MAR-2001.  
XX  
PF 27-JAN-2000; 2000DE-02001395.  
XX  
PR 27-JAN-2000; 2000DE-02001395.  
XX  
PA (GPCB-) GPC BIOTECH AG.  
XX  
DR MPI; 2002-228394/29.  
DR P-PSDB; ABB08834.  
XX  
PT New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
PS Claim 1; Fig 1; 19pp; German.  
XX  
CC The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio  
CC

XX  
SQ Sequence 678 BP, 198 A; 147 C; 159 G; 174 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922A-12 (1-225) x ABA95905 (1-678)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGATCTTCTAAGAACGTTCATCAAGGAATTCAGATTCAGAGTTAGAAAGTGAAGT 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACTGTTAACGGTCAACGAATTCGAAATCGAAGGTGAAGTGAAGTGAACCATACGAAGT 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 121 CACAACACTGTCAGTTGAAGGTTACTTAAGGGTGGTCCATTGCCATTGGGACATC 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGTCTCCACAATTCCAATAACGTTCTTAAGGTCTACGTCAAAGCACCAGCTGACATTCCA 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 241 GACTACAGAAGTGTTCCTTCCAGAGAGTTTCAAGTGGGAAAGATCATGAACCTTCGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGTGTGTACTGTACTCAAGACTCCCTCTTCAAGACGGTGTCTCACTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140  
Db 361 AAGGTCAAGTTCAATCGGTGTCAACTTCCCATCTGACGGTCCAGTCATGCCAAAGAAGACT 420  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGTGGGAAGCTTCTACCGAAGCTTTGTATACCAAGACGGTGTCTGAAGGTGA 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGCGCTTGAAGTTGAAGACGGTGTCACTTGGTGAATTCAAGTCTATC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCTAAGAAGCCAGTCCAAATGCCAGGTACTACTAGTTGACTCTTAAGTTGAC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCTCACAACGAAGACTACACTATCGTCGAACAATACGAACGTACTGAAGGTAGA 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTTGTCTTGTG 675  
RESULT 5  
ID ABA95922 standard; DNA; 678 BP.  
XX ABA95922;  
AC ABA95922;  
XX 29-MAY-2002 (first entry)  
DT 29-MAY-2002 (first entry)  
DE Yeast optimised RFP related DNA SEQ ID NO 19.  
XX  
KW Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW



KW Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
XX Anthozoa.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "yeast optimised Red Fluorescent Protein"  
XX  
PN DE20001395-U1.  
XX  
PD 15-MAR-2001.  
XX  
XX 27-JAN-2000; 2000DE-02001395.  
PF 27-JAN-2000; 2000DE-02001395.  
XX  
PR 27-JAN-2000; 2000DE-02001395.  
XX  
XX (GPCB-) GPC BIOTECH AG.  
XX  
XX WPI; 2002-228394/29.  
DR  
XX  
XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
XX Disclosure; Page 14-15; 19pp; German.  
XX  
XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
CC useful to the invention  
XX  
XX SQ Sequence 678 BP; 202 A; 118 C; 159 G; 199 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservat: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922a-12 (1-225) x ABA95922 (1-678)  
QY 1 MetArgSerSerLySAenValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGATCTTCTTAAGAACGTTATTAAAGCAATTCATGAGATTCAAAGCTTAGAATGGAAGT 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACTGTTAAACGGTCACGAATTCGAAATTGAAGGTGAAGGTGAAGGTAGACCATAACGAAGT 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAACACTGTTAAGTTGAAGGTTAAAGGTGTCATTGCCATTGCTTGGAACATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 181 TTGTCTCCACAATCCAAATACGGTCTAAAGTTTACGTTAAGCACCCAGCTGACATTCCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACAGAAGTGTCTTCCCAAGAGTTTCAAGTGGGAAAGAGTTATGAACCTCGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGTGTGTGTACTGTACTCAAGACTCTTCTTGCAAGACGGTGTTCATTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTTAAGTTCAATGTTGTTAACTTCCATCTGACGGTCCAGTTATGCAAAAGAGACT 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGTTGGGAAGCTTCTACTGAAAGATTGTACCCAAGAGACGGTGTGTAAGGGTGAA 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
DB 481 ATTCACAGAGCTTTGAAGTTGAAGACGGTGTCACTACTTGGTTGAATTCAAGTCTATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCTAAGAAGCCAGTTCATTGCCAGGTTACTACTAGCTGAAGTTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATTACTTCTCAACAAGAGACTACTATTGTTGAACATAACGAAGACTGAAGTAGA 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACCTGTCTTG 675  
RESULT 6  
ABA95921  
ID ABA95921 standard; DNA; 678 BP.  
XX  
XX ABA95921;  
AC  
XX  
DT 29-MAY-2002 (first entry)  
XX  
DE Yeast optimised RFP related DNA SEQ ID NO 18.  
XX  
XX Yeast; red fluorescent protein; RFP; plant; transgenic; GRP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
XX Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
XX  
OS Anthozoa.  
XX  
XX Key Location/Qualifiers  
FH CDS 1..678  
FT /\*tag= a  
FT /product= "yeast optimised Red Fluorescent Protein"  
XX  
PN DE20001395-U1.  
XX  
PD 15-MAR-2001.  
XX  
XX 27-JAN-2000; 2000DE-02001395.  
PF 27-JAN-2000; 2000DE-02001395.  
PR 27-JAN-2000; 2000DE-02001395.  
XX  
XX (GPCB-) GPC BIOTECH AG.  
XX  
XX WPI; 2002-228394/29.  
DR P-PSDB; ABB08834.  
XX  
XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
XX Disclosure; Page 14; 19pp; German.  
XX  
XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or



CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
CC useful to the invention  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABA95921 (1-678)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAGATGTATCAAGAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAAGGCGAAGAGAGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGACCTTGCCATTGTGGGATATTT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGACCCCTGCCACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCTATTCTCGAAGGATTTAATGGGAAAGGTCATGAACCTTGA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTGTACTGTAAACCAAGATTCAGTTGCGAGATGGCTGTTCACTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCCCTCCGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCAGCAGCGTTTGTATCCTCGTATGGCGGTGTGAAGAGAGAG 480  
QY 161 IleHisLeuAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAAAGAACCTGTGCAGCTACCAAGGTAATACTGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGlyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTCTT 675

RESULT 7  
ABA95920  
ID ABA95920 standard; RNA; 678 BP.  
XX

AC ABA95920;  
XX  
XX 29-MAY-2002 (first entry)  
DT  
XX  
XX Yeast optimised RFP encoding RNA SEQ ID NO 16.  
DE  
XX  
XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology; ss.  
XX  
XX Anthozoa.  
OS

XX DE20001395-U1.  
XX  
XX 15-MAR-2001.  
PD  
XX  
XX 27-JAN-2000; 2000DE-02001395.  
PF  
XX  
XX 27-JAN-2000; 2000DE-02001395.  
PR  
XX  
XX 27-JAN-2000; 2000DE-02001395.  
PA  
XX (GPCB-) GPC BIOTECH AG.  
XX WPI; 2002-228394/29.

PT New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
PS

Disclousre; Page 13; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of an RNA sequence corresponding to the yeast optimised  
CC RFP encoding cDNA given as SEQ ID NO 1 (ABA95905)  
XX

SQ Sequence 678 BP; 198 A; 147 C; 159 G; 0 T; 174 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABA95920 (1-678)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 AUGAGAUUCUUAAGACGUCAUCCAAGGAUUCAGAUUCCAAGGUVAGAUUGGAAGGU 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACUGUUAACGGUCACGAUUCGAUAUCGAAGUGAAGUGAAGUAGACCAUACGAAGU 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAAACACUGUCAAGUUGAAGGUVUACAAGGUGUCCAUUGCCAUUGCGTUGGACAUC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 UUGUCUCCACAUCUCCAUCUACGUGUUCUAAAGUCUACGUAAGCACCCAGCUCAGAUUCCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100



|||||  
Db 241 GACUACAAGAGUTGUCUCCAGAGAGUUCAGUGGGAAGAGUCUAGACUUCGAA 300  
Qy 101 AaPGLyGLyValThrValThrGlnAaPSeSerLeuGlnAaPGLyCySPheIleTyr 120  
Db 301 GACGGUGUGUGUUAUGUUAUCUACAAGACUCCUUGCAAAGACGGUGUUAUCUAC 360  
Qy 121 LysValLysPheIleGLyValAaSPheProSeRaSPGLyProValMetGlnLysLysThr 140  
Db 361 AAGGUCAAAGUUCAGGUGUACAACUCCCAUUGACGGUCCAGUCAGUCAAAGAAAGACU 420  
Qy 141 MetGLyTrpGluAlaSerThrGluArgLeuTyrProArgAaSPGLyValLeuLysGLyGlu 160  
Db 421 AUGGGUGGGAGGUCUACCGAACGUGUUGACCAAGAGACGGUGUCCUUAAGGUGAA 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAaSPGLyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 AUCCACAAGGCCUUGAAGUUGAAGGACGGUGUCAUAUUGGUGCAAUAUCAAAGUCUAUC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGLyTyrTyrTyrValAaSPSerLysLeuAaSP 200  
Db 541 UACAUGGCUAAGAGCCAGUCCAAUUGCCAGGUAUAUAUAUAUAUAUAUAUAUAUAUAUA 600  
Qy 201 IleThrSerHisAaNGluAaSPTyrrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 AUCACCUUCUACAACGAAGACUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAUA 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACUUGUUCUUG 675

RESULT 8  
AAD46278  
ID AAD46278 standard; DNA; 678 BP.

XX AAD46278;

XX 27-DEC-2002 (first entry)

XX Discosoma sp. drFP583 (NFP-6) wild-type protein encoding DNA.

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
XX fluorescence activated cell sorting application; fluorescent timer;  
XX biosensor; fluorescence resonance energy transfer application; FRRT;  
XX colouring agent; recombinant DNA application; analyte detection assay;  
XX sunscreen; second messenger detector; drFP583 protein; NFP-6; gene; ds.

XX Discosoma sp.

XX Key Location/Qualifiers

FT CDS 1..678  
FT /\*tag= a  
FT /product= "drFP583 wild-type protein"

XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX 21-FEB-2001; 2001US-0270983P.

XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;

XX WPI; 2002-691654/74.

XX P-PSDB; AAE28833.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting

PT applications.  
XX  
PS Disclosure; Page 70; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is a DNA encoding Discosoma sp. drFP583 (NFP-6) wild-type  
XX protein of the invention

SQ Sequence 678 BP, 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AAD46278 (1-678)

Qy 1 MetArgSerSerLysAaSPValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 1 ATGCGCTCCTCCAGAAGAGCTCATCAAGAGCTTCAATGCGCTTCAAGGTGCGCATGAGGGC 60

Qy 21 ThrValAaNGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40

Db 61 ACCGTGAACGGCCAGAGTTGAGATCGAGGCGGAGGGCGGCGCCCTTCAAGGATCGAGGGC 120

Qy 41 HisAaThrValLysLysLeuLysValThrLysGLyGlyProLeuProPheAlaTrpAaSPile 60

Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGCCCTGCTTGCCTGGGACATC 180

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAaSPilePro 80

Db 181 CTGTCCCCCAGTTCAGTAGCGGCTCCAAGGTAGTGAAGCACCCGCCGACATCCCC 240

Qy 81 AaPTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAaSPheGlu 100

Db 241 GACTACAAGAGCTGTCTTCCCGGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCGAG 300

Qy 101 AaPGLyGLyValValThrValThrGlnAaPSeSerLeuGlnAaPGLyCySPheIleTyr 120

Db 301 GACGGCGGCGTGTGACCTGACCCAGAGACTCTCTCTGCAAGACGGCTCTCATCTAC 360

Qy 121 LysValLysPheIleGLyValAaSPheProSeRaSPGLyProValMetGlnLysLysThr 140

Db 361 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAAGACC 420

Qy 141 MetGLyTrpGluAlaSerThrGluArgLeuTyrProArgAaSPGLyValLeuLysGLyGlu 160

Db 421 ATGGGCTGGAGGCGCTCAACGAGCGCCTGTACCCCGCAGCGCGTGTGAAGGCGAG 480

Qy 161 IleHisLysAlaLeuLysLeuLysAaSPGLyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATCCAACAAGGCCCTGAAGCTGAAGAGACGGCGGCACTAAGTGTGAAGTCAAGTCCATC 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGLyTyrTyrTyrValAaSPSerLysLeuAaSP 200

Db 541 TACATGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAGCTGGAC 600

Qy 201 IleThrSerHisAaNGluAaSPTyrrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATCACTCCCAACAAGAGACTACACCATCGTGAAGAGTACGAGCGCACCGAGGCGCCG 660

Qy 221 HisHisLeuPheLeu 225



Db 661 CACCACCTGTTCTG 675

RESULT 9

AAD28207

ID AAD28207 standard; DNA; 678 BP.

XX

AC AAD28207;

XX

DT 22-APR-2002 (first entry)

XX

DE Discosoma sp. humanised wild-type Anthozoa protein drFP583 DNA.

XX

KM Fluorescent timer protein; protein movement; translocation; trafficking;

KM promoter activity; gene expression; transgenic plant; gene modification;

KM protein age; anthozoa protein; drFP583; ds.

XX

OS Discosoma sp.

XX

FH Key location/Qualifiers

FT CDS 1..678

FT /\*tag= a

FT /product= "Humanised wild-type Anthozoa protein drFP583"

XX

PN W0200196373-A2.

XX

PD 20-DEC-2001.

XX

PF 13-JUN-2001; 2001WO-US019097.

XX

PR 14-JUN-2000; 2000US-0211607P.

XX

PA (CLON-) CLONTECH LAB INC.

XX

PI Fradkov AF, Terakikh A;

XX

XX WPI; 2002-154595/20.

DR P-PSDB; AAE17540.

XX

PT New fluorescent timer proteins comprising an emission spectrum that

PT changes over time from a first wavelength to a second wavelength, useful

PT for monitoring intracellular protein movement, translocation, trafficking

PT or stability.

XX

XX Example 1; Fig 1; 89pp; English.

XX

CC The invention relates to a fluorescent timer protein having an emission

CC spectrum that changes over time after synthesis from a first wavelength

CC to a second wavelength. The fluorescent timer proteins are useful in

CC monitoring the activity of a promoter, determining the age of a protein,

CC identifying an agent that modulates the activity of a promoter and in

CC enriching a population of cells comprising a fluorescent timer protein.

CC The fluorescent timer proteins are also useful for assessing gene

CC expression during development of a multicellular organism or during

CC cellular differentiation, in response to a drug or other inducer of

CC promoter activity, as a reporter to serve as a read-out of promoter

CC activity, monitoring intracellular protein movement or translocation,

CC protein trafficking, or protein stability, to investigate temporal

CC aspects of the activity of a regulatory element, for determining cell

CC fate during development and organ remodelling, in spatial and temporal

CC visualisation of newly synthesised proteins and accumulated proteins, and

CC in distinguishing between newly formed and pre-existing structures, e.g.

CC membrane junctions and extracellular matrix components. The fluorescent

CC timer proteins may further be used to investigations where photobleaching

CC techniques are employed, as detectable labels, as selectable markers, as

CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage

CC assays, and as second messenger detectors. The nucleic acids can be used

CC to generate transgenic, non-human plants or animals or site-specific gene

CC modifications in cell lines. The present sequence is a DNA encoding

CC Discosoma sp. humanised wild-type Anthozoa protein drFP583 used for

CC generating fluorescent proteins

XX

XX Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 2.3e-140

Score: 1214.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 6

US-10-006-922A-12 (1-225) x AAD28207 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20

Db 1 ATGCGCTCTCCAGAACGTCAATCAAGAGTTCTATGCGCTTCAAGGTGCGCATGAGGGC 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120

QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60

Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTTCCCTGCGCTGGACATC 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80

Db 181 CTGTCCCCCAGTTCCAGTACGAGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 240

QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100

Db 241 GACTACAGAAGAGCTGTCTTCCCCGAGGCTTCAAGTGGAGCGCGTGATGACTTCGAG 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 301 GACGGCGCGTGTGACCGTGACCCAGGACTCTCTCCCTGCAGGACGGCTGCTTACTAC 360

QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140

Db 361 AAGGTGAAGTTATCGGCGTGAACCTTCCCCCTCCAGCGGCCCGTGTATGCAAGAGAACCC 420

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlu 160

Db 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCGACGGCGTGTAAAGGGCAG 480

QY 161 IleHisLySAlaLeuLySLeuLySAspGlyHisTyrLeuValGluPheLySerIle 180

Db 481 ATCCACAAGGCCCTGAAGCTGAAGAGCGCGGCCACTACTGTGAGTTCAAGTCCATC 540

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 541 TACATGGCCAGAAGCCCGTGCAAGCTGCCCGGCTACTACTAGCTGACTCCAAGCTGAC 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATCACTCCACACAGAGACTACACCATCTGTGAGCAGTACGAGCGACCGAGGGCCGC 660

QY 221 HisHisLeuPheLeu 225

Db 661 CACCACCTGTTCTG 675

RESULT 10

ADC24127

ID ADC24127 standard; DNA; 678 BP.

XX

AC ADC24127;

XX

DT 18-DEC-2003 (first entry)

XX

DE Discosoma wild-type red fluorescent protein DNA #1.

XX

KM Discosoma red fluorescent protein; DsRed; AB interface; AC interface;

KM fluorescent protein variant; transcription induction detection;

KM fluorescence energy resonance transfer; FRRT; protein kinase;

XX

XX protein phosphatase; ion indicator; ds.



OS Discosoma.  
XX  
PN US2003059835-A1.  
XX  
PD 27-MAR-2003.  
XX  
PF 10-APR-2002; 2002US-00121258.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
XX  
PI Tsien RY, Campbell RE;  
XX  
DR WPI, 2003-743764/70.  
DR P-PSDB; ADC24126.  
XX  
PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
XX  
PS Example 1; SEQ ID NO 2; 67bp; English.  
XX  
CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC the spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma wild-  
CC type red fluorescent protein.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922A-12 (1-225) x ADC24127 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
|||||  
Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGAGTTTCATGAGTTTAAAGTTCCGATGGAAGA 60  
|||||  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||||

Db 61 ACGGTCATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLyValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||  
Db 121 CACATAACCGTAAGCTTAAGCTTAACCAAGGGGGACCTTGCCATTGCTTGGAATAT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLysHisProAlaAspIlePro 80  
|||||  
Db 181 TTGTCAACCAATTTCAGTATGGAACCAAGGTATATGTCAAGCACCCTGCCACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||||  
Db 241 GACTATTAAGAAAGCTGTCAATTTCTGAAGCAATTTAAATGGGAAAGGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||  
Db 301 GACGGTGGCGTCGTTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTCAITCAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||||  
Db 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCGATGAGACCTGTATGCAAAAGAAAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
|||||  
Db 421 ATGGGCTGGGAAGCCACGACACTGACGTTTGATCTCCTGATGGCGTGTGAAGAGAGAG 480  
QY 161 IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||||  
Db 481 ATTCAATAGGCTCTGAAGCTGAAGACGGTGTTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||||  
Db 541 TACATGGCAAGAAGCCTGTGCAGTACCAAGGTACTACTAATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||  
Db 601 ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAAGAACCGAGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
|||||  
Db 661 CACCATCTGTTCTT 675

RESULT 11  
ADF70404  
ID ADF70404 standard; cDNA; 678 BP.  
AC ADF70404;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Discosoma wild-type GFP variant cDNA SeqID27.  
XX  
KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
KW GFPuv; Enhanced GFP; EGFP; gene; ss.  
XX  
OS Discosoma sp.  
XX  
PN WO2003071272-A1.  
XX  
PD 28-AUG-2003.  
XX  
PF 21-FEB-2003; 2003WO-JP001901.  
XX  
PR 22-FEB-2002; 2002JP-00045728.  
PR 23-JUL-2002; 2002JP-00213949.  
PR 11-OCT-2002; 2002JP-00298237.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hasegawa M;  
XX  
DR WPI; 2003-697654/66.



DR P-PSDB; ADF70403.  
XX Transformation of cells with a fusion protein of an orphan receptor  
PT protein with a fluorescent protein useful for identification of ligands  
PT to the orphan receptor.  
XX  
PS Disclosure; SEQ ID NO 27; 594pp; Japanese.  
XX  
CC This invention relates to a novel method of identifying ligands to an  
CC orphan receptor protein which comprises transforming cells with DNA  
CC encoding a fusion protein of the orphan receptor with a fluorescent  
CC protein, so that the fusion protein is expressed in the cells (or cell  
CC membranes isolated from them) and contacting the cells with the potential  
CC ligand to be tested. A suitable fluorescent protein (GFP), for example GFP-1,  
CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
CC identification of ligands binding to an orphan receptor protein.  
XX  
SQ Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADF70404 (1-678)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAAGATGTTATCAAGAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGGAGAGGGAGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGACCTTGCCATTGGCTGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCAAGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGCGTGGGAAGCCAGCACTGACGCTTGTATCCTCGTATGGCCGTGTGAAGAAGAGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGACGCGTGTCAATTAAGTTGAATTCAAAGATATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGlyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAAGAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGGCC 660

QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTCTT 675

RESULT 12  
ID ADL46204 standard; DNA, 678 BP.  
XX  
AC ADL46204;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma red fluorescent protein (DsRed) coding sequence.  
XX  
KW ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.

Key Location/Qualifiers  
FH 1.678  
FT CDS /\*tag= a  
FT /product= "DsRed protein"

WO2003086446-A1.

23-OCT-2003.

09-APR-2003; 2003WO-US010879.

10-APR-2002; 2002US-00121258.  
29-JUL-2002; 2002US-00209208.

(REBC ) UNIV CALIFORNIA.

Tsien RY, Campbell RE, Baird GS;

WPI; 2003-845265/78.  
P-PSDB; ADL46203.

New monomeric and dimeric Anthozoan fluorescent protein variants with reduced propensity to oligomerize, and encoding polynucleotides, useful in molecular biology, e.g. in immunoassays or in tracking protein movement in cells.

Disclosure; SEQ ID NO 2; 166pp; English.

The invention relates to a polynucleotide sequence encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize. The protein variant comprises one or more amino acid substitutions at the AB and/or AC interface(s) of the wild-type DsRed sequence, where the substitutions result in reduced propensity of the DsRed variant to form tetramers and where the variant displays detectable fluorescence of at least one red wavelength. The composition and methods are useful in producing red fluorescent proteins having reduced propensity for oligomerization, especially tetramerization. The protein may be used in molecular biology and in other scientific applications, such as in immunoassays or hybridization assays, or in tracking the movement of proteins in cells. This sequence corresponds to the DsRed coding sequence.

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46204 (1-678)



QY

1

MetArgSerSerLysAsnValIleIleGluPheMetArgPheLysValArgMetGluGly

20

DB

1

ATGAGGCTTCCAGAATGTTATCAAGAGCTTCATGAGGTTTAAGTTCCGATGGAAGA

60

QY

21

ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly

40

DB

61

ACGGTCAATGGGCAAGATTGTGAATAGAAAGCGGAAGAGGGGAGGCCATACGAAGGC

120

QY

41

HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle

60

DB

121

CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGAATATT

180

QY

61

LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro

80

DB

181

TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA

240

QY

81

AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu

100

DB

241

GACTATAAAAAGCTGTCATTCTCGAAGGATTTAAATGGGAAGGTCATGAACCTTGAA

300

QY

101

AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

DB

301

GACGGTGGCGTCTTAAGTCTTAACCCAGGATTCAGTTTGCAAGATGGCTTTTCATTAC

360

QY

121

LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr

140

DB

361

AAGTCAAGTTCATTGCGCGTGAACCTTCCGATGACCTGTTATGCAAAAGAAGACA

420

QY

141

MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

DB

421

ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCCTCGTATGGCGGTGTAAGAAGAG

480

QY

161

IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle

180

DB

481

ATTCATAAGGCTCTGAGCTGAAGAAGCGGTGTCATTACTGAATGTAATCAAAAGTATT

540

QY

181

TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp

200

DB

541

TACATGGCAAGAAGCCTGTGCAGCTACCAAGGGTACTACTATGTGACTCCAACTGGAT

600

QY

201

IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg

220

DB

601

ATAACAAGCCACAAACGAAGACTATACAAATCGTTGAGCAGTATGAAGAACCAGGGAAGCC

660

QY

221

HisHisLeuPheLeu

225

DB

661

CACCATCTGTCTCTT

675

RESULT 13

ADN33978

ADN33978 standard; DNA; 678 BP.

XX

ADN33978;

XX

01-JUL-2004 (first entry)

XX

wild-type DSRED encoding sequence.

DE

Cnidarian, fluorescence resonance energy transfer; FRET; wild-type DSRED;

KW

ds.

XX

DibcoSoma sp.

OS

WO2003054158-A2.

XX

03-JUL-2003.

PD

18-DEC-2002; 2002WO-US040539.

XX

19-DEC-2001; 2001US-0341723P.

PR

(UYCH-) UNIV CHICAGO.

XX

PA

XX

PI

Bevis B, Glick B;

XX

DR

WPI; 2003-569236/53.

DR

P-PSDB; ADN33979.

XX

PT

Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent

PT

mutant of a Cnidarian chromo- or fluorescent protein or its mutant,

PT

useful for applications involving chromo- or fluorescent proteins.

XX

PS

Claim 8; SEQ ID NO 1; 65bp; English.

XX

CC

The present invention relates to nucleic acid that encodes a rapidly

CC

maturing chromo or fluorescent mutant of a Cnidarian chromo- or

CC

fluorescent protein or its mutant. The protein is useful in applications

CC

involving nucleic acid encoding a chromo- or fluorescent protein and is

CC

useful for producing a chromo and/or fluorescent protein which involves

CC

growing the cell, whereby the protein is expressed, and isolating the

CC

protein substantially free of other proteins. The protein is useful in

CC

applications involving chromo- or fluorescent protein and is useful as

CC

PCR primers, hybridization probes, etc. The expression cassettes are

CC

useful for synthesizing related proteins. The chromoproteins are useful

CC

as coloring agents which are capable of imparting color or pigment to a

CC

particular composition of matter e.g. food compositions, pharmaceuticals,

CC

cosmetics, living organisms, e.g., animals and plants. The chromoproteins

CC

may also find use as labels in analyte detection assays, e.g. assays for

CC

biological analytes of interest and as selectable markers in recombinant

CC

DNA applications, e.g. the production of transgenic cells and organisms.

CC

The fluorescent proteins find use in a variety of different applications,

CC

e.g. in fluorescence resonance energy transfer (FRET) applications, as

CC

biosensors in prokaryotic and eukaryotic cells, in applications involving

CC

the automated screening of arrays of cells expressing fluorescent

CC

reporting groups by using microscopic imaging and electronic analysis, as

CC

second messenger detectors, and in fluorescence activated cell sorting

CC

applications and as in vivo marker in animals. The fluorescent proteins

CC

also find use in protease cleavage assays. The proteins can also be used

CC

is assays to determine the phospholipid composition in biological

CC

membranes and as a fluorescent timer. The present sequence represents the

CC

wild-type DSRED encoding sequence.

XX

SEQ

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

2.3e-140

Length:

678

Score:

1214.00

Matches:

225

Percent Similarity:

100.00%

Conservative:

0

Best Local Similarity:

100.00%

Mismatches:

0

Query Match:

100.00%

Indels:

0

DB:

11

Gaps:

0

US-10-006-922A-12 (1-225) x ADN33978 (1-678)

QY

1

MetArgSerSerLysAsnValIleIleGluPheMetArgPheLysValArgMetGluGly

20

DB

1

ATGAGGCTTCCAGAATGTTATCAAGAGCTTCATGAGGTTTAAGGTTCCGATGGAAGA

60

QY

21

ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly

40

DB

61

ACGGTCAATGGGCAAGATTGTGAATAGAAAGCGGAAGAGGGGAGGCCATACGAAGGC

120

QY

41

HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle

60

DB

121

CACAATACCGTAAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGAATATT

180

QY

61

LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro

80

DB

181

TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA

240

QY

81

AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu

100

DB

241

GACTATAAAAAGCTGTCATTCTCGAAGGATTTAAATGGGAAGGTCATGAACCTTGAA

300

QY

101

AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120



|||||  
Db 301 GACGGTGGCGTCTGTTACTGTATACCCAGGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAAGACA 420  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCGTGAATGGCGTGTGAAAGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCGTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGAGT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGATGAAAGAACCAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCCT 675

RESULT 14  
AD136420  
ID AD136420 standard; DNA; 678 BP.

AC AD136420;  
XX  
XX 22-APR-2004 (first entry)  
DT

XX Discosoma sp. red fluorescent protein (RED) DNA.

XX Fluorobody; binding ligand; green fluorescent protein; GFP;  
KM target detection; gene; red fluorescent protein; ds; RED.

XX Discosoma sp.

XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "Discosoma sp. red fluorescent protein (RED)"

XX US2003203355-A1.

XX 30-OCT-2003.

XX 24-APR-2002; 2002US-00132067.

XX 24-APR-2002; 2002US-00132067.

XX (LALA-) LOS ALAMOS NAT LAB.

XX (REGC ) UNIV CALIFORNIA.

XX Bradbury AM, Zeytun A, Waldo GS;

XX WPI; 2004-154325/15.

XX P-PSDB; AD136421.

XX Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.

XX Example 6; SEQ ID NO 3; 23pp; English.

XX The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The

CC present sequence is Discosoma sp. red fluorescent protein (RED) DNA used  
CC in the exemplification of the invention.

XX SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-006-922a-12 (1-225) x AD136420 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyGly 20  
Db 1 ATGAGGCTCTCCAGAAGATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGlyGly 40  
Db 61 ACGGTCAATGGGCACGACTTTGAATATGAAGCCGAAGAGAGAGAGGCCATACGAAGCC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGCTTGGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTCAATGATGGAAGCAAGGTATATGTCAAAGCACCTTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGlyGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAGCTGTCATTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACTTTGA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTGTTACTGTAAACCAAGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGAACCTGTTATGCMAAAGAAGACA 420

QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCGTATGGCGGTGTGAAGAAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCGTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGAGT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACGAGGAGCGC 660

QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCCT 675

RESULT 15  
ADM97768  
ID ADM97768 standard; DNA; 678 BP.

XX ADM97768;

XX 01-JUL-2004 (first entry)  
DT

XX D sp red fluorescent protein coding sequence SEQ ID NO: 21.



KW ds; gene; enzyme; sensor cell; fluorescent protein;  
KW signal transduction detection system; promoter; targeting sequence;  
KW targeted drug.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "fluorescent protein"  
XX  
PN WO2004031415-A2.  
XX PD 15-APR-2004.  
XX PF 05-SEP-2003; 2003WO-US028078.  
XX PR 05-SEP-2002; 2002US-0408297P.  
XX PA (VERT-) VERTEX PHARM INC.  
XX PI Whitney MA, Zeh K, Sanders PS;  
XX DR WPI; 2004-330208/30.  
XX P-PSDB; ADM97769.  
XX  
PT Developing a sensor cell, useful in determining the activity of a target  
PT gene and in developing therapeutic drugs, comprises providing cells  
PT comprising a signal transduction detection system and introducing DNA  
PT construct into cells.  
XX  
PS Disclosure; Page 167-168; 234pp; English.  
XX  
CC The present invention relates to a method of developing a sensor cell,  
CC for determining the activity of a target gene in the cell, which  
CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing  
CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
CC coding sequence shown in the exemplification of the invention.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatch: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-10-006-922A-12 (1-225) x ADM97768 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAAGATGTTATCAAGAGTTCATGAGTTTAAAGTTCCGATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCAAGATTGAAATAGAAGCGGAAGGAGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAAGTTAACCAAGGGGGGACCTTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTyrGluArgValMetAsnPheGlu 100

Db 241 GACTATAAAAAGCTGTCAATTCTGAAAGATTAAATGGGAAAGGTCATGAACTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTGTACTGTAAACCAAGATTCCAGTTTGACAGATGGCTGTTCATCTAC 360  
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140  
Db 361 AAGGTCAAGTTCAATTGGCGTGAACCTTCCCTCCGATGGACCTGTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGACCGTTTGATCCTCGTGAATGGCGGTGGAAGAGAG 480  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
Db 481 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
Db 541 TACATGCAAGAAGACCTGTGCACTAACAGGTAAGTACTATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCTT 675

Search completed: January 12, 2006, 10:16:26  
Job time : 485 secs



**THIS PAGE BLANK (USPTO)**



Gencore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 09:14:42 ; Search time 3744 Seconds  
(without alignments)  
2811.724 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO.spool.p/US10006922/runat.10012006.162408.18933/app\_query.fasta\_1.391  
-DB=EST\_QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922@CGN.1.1.8010@runat.10012006.162408.18933 -NCPU=6 -ICPU=3  
-NO\_MMWP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hlc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_g881:\*  
10: gb\_g882:\*  
11: gb\_g883:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	461	38.0	556	8 DR987865	DR987865 JGI_AOSF1
2	461	38.0	556	8 DR987866	DR987866 JGI_AOSF1
3	225	18.5	661	3 BI385436	BI385436 BFL26_000
4	218.5	18.0	508	5 BW802012	BW802012 BW802012
5	218.5	18.0	646	5 BW736869	BW736869 BW736869
6	217.5	17.9	639	5 BW736108	BW736108 BW736108
7	216.5	17.8	604	5 BW738638	BW738638 BW738638

8	216.5	17.8	606	5 BW741633	BW741633 BW741633
9	216.5	17.8	637	3 BI379468	BI379468 BFLG1_001
10	216.5	17.8	687	3 BI387917	BI387917 BFL26_002
11	212.5	17.5	643	3 BI386699	BI386699 BFL26_001
12	212.5	17.5	675	3 BI382638	BI382638 BFLG2_000
13	210.5	17.3	590	5 BW883955	BW883955 BW883955
14	210.5	17.3	812	5 BW893733	BW893733 BW893733
15	209.5	17.3	790	1 AL044652	AL044652 DKFZP434P
16	206.5	17.0	630	5 BW731527	BW731527 BW731527
17	206.5	17.0	677	5 BW894392	BW894392 BW894392
18	205.5	16.9	605	5 BW744973	BW744973 BW744973
19	205.5	16.9	618	5 BW871203	BW871203 BW871203
20	205.5	16.9	668	5 BW894220	BW894220 BW894220
21	203.5	16.8	555	5 BW845436	BW845436 BW845436
22	203.5	16.8	602	5 BW872238	BW872238 BW872238
23	203.5	16.8	620	5 BW870345	BW870345 BW870345
24	203.5	16.8	630	5 BW740938	BW740938 BW740938
25	203.5	16.8	641	5 BW798018	BW798018 BW798018
26	203.5	16.8	641	5 BW881726	BW881726 BW881726
27	203.5	16.8	647	5 BW736374	BW736374 BW736374
28	203.5	16.8	652	5 BW733673	BW733673 BW733673
29	203.5	16.8	652	5 BW81040	BW81040 BW81040
30	202.5	16.7	602	5 BW870882	BW870882 BW870882
31	202.5	16.7	744	8 DR622961	DR622961 EST101308
32	201.5	16.6	573	5 BW865453	BW865453 BW865453
33	201.5	16.6	675	5 BW895120	BW895120 BW895120
34	200.5	16.5	722	8 DR624559	DR624559 EST101468
35	200.5	16.5	730	8 DR623458	DR623458 EST101358
36	200.5	16.5	736	8 DR629310	DR629310 EST101943
37	200.5	16.5	768	8 DR629433	DR629433 EST101956
38	200.5	16.5	791	8 DR621827	DR621827 EST101195
39	200.5	16.5	795	8 DR624841	DR624841 EST101496
40	200.5	16.5	818	8 DR627300	DR627300 EST101742
41	200.5	16.5	828	8 DR622639	DR622639 EST101276
42	200.5	16.5	831	8 DR628628	DR628628 EST101875
43	200.5	16.5	833	8 DR624844	DR624844 EST101497
44	200.5	16.5	840	8 DR624776	DR624776 EST101490
45	200.5	16.5	862	8 DR624009	DR624009 EST101413

ALIGNMENTS

RESULT 1  
DR987865/c 556 bp mRNA 1linear EST 03-AUG-2005  
LOCUS JGI\_AOSF1131.rev AOSF Montastreae faveolata adult colony  
DEFINITION Montastreae faveolata cDNA clone AOSF1131 3', mRNA sequence.

ACCESSION DR987865 GI:71782571  
VERSION DR987865  
KEYWORDS EST.  
SOURCE Montastreae faveolata  
ORGANISM Montastreae faveolata  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
Favilina; Favillidae; Montastreae.  
REFERENCE 1 (bases 1 to 556)  
AUTHORS Schwarz,J.A., Brokstein,P., Manohar,C., Coffroth,M.A., Szmant,A.  
and Medina,M.  
TITLE Coral-Symbiodinium EST Project  
JOURNAL Unpublished (2005)  
COMMENT Other\_ESTs: JGI\_AOSF1131.fwd  
Contact: Schwarz, JA, Medina, M.  
Evolutionary Genomics  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925-296-5823  
Email: jaschwarz@lbl.gov  
cDNA Library Preparation: DOE Joint Genome Institute:  
http://www.jgi.doe.gov  
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  
Naming Conventions: EST name is generated by the concatenation of  
the JGI Clone Id and the direction of sequencing. The suffix '.rev'  
indicates a reverse sequencing read of the insert. It does not  
necessarily reflect the orientation of the insert.



Poly-T: A run of 14 or more T residues at the beginning of this sequence has been removed.  
Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.  
Plate: AOSF 0009 row: f column: 20  
High quality sequence stop: 556.

FEATURES  
Source  
Location/Qualifiers

1.556  
/organism="Montastraea faveolata"  
/mol\_type="mRNA"  
/db\_xref="taxon:48498"  
/clone="AOSF1131"  
/dev\_stage="Adult colony"  
/lab\_host="Electromax DH10B"  
/clone\_lib="AOSF Montastraea faveolata adult colony"  
/note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; The library was prepared from total RNA using the Creator SMART cDNA Library Construction Kit with the LD-PCR method to amplify the cDNA. Amplified cDNA was digested with SfiI, size selected for >400bp, and ligated into the pDNR-LIB vector. WARNING: this library contains a small percentage of cDNAs derived from the coral's symbiont, Symbiodinium sp."

ORIGIN

Alignment Scores:

Pred. No.: 3.93e-45 Length: 556  
Score: 461.00 Matches: 83  
Percent Similarity: 62.96% Conservative: 36  
Best Local Similarity: 43.92% Mismatches: 46  
Query Match: 37.97% Indels: 24  
DB: 8 Gaps: 1

US-10-006-922A-12 (1-225) x DR987865 (1-556)

QY 32 GUGUGUGUGUAGArgProTyrGUGUGUHisAAnthrValLysLeuLysValThrLysGly 51  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 556 GACGGAGAGGCAAGCCTTTCGAGGAGAAACACAGATATAACCTTCAAGTCCAGAAAGGT 497  
QY 52 GlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysVal 71  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 496 GGAACCTCTCCCTTTGCTTACGATATCTTGACGACAGCATTTCTCGTACGGCAACAGGCA 437  
QY 72 TyrValLysHisProAlaAspIleProAspTyrLysLysLeuSerPheProGlnGlyPhe 91  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 436 TTCACCAATAACCCAGAGACCTCCCAAGACTATTTCAGCAGATCGTTCTCCGGGGTAT 377  
QY 92 LysTrpGlnArgValMetAsnPhelGluAspGlyGlyValValThrValThrGlnAspSer 111  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 376 TCCTGGGAACGTTGTATGACTTTCCGAAGATGAGGCATTGGCACCGTGTCAACAACATC 317  
QY 112 SerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheproSer 131  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 316 AAAATGAGGCGACCGTTTATCTATGATATTCGATTTCATGTTGTTAACTTTCCTGCT 257  
QY 132 AspGlyProValMetGlnLysLysThrMetGlyTyrGlnAlaSerThrGluArgLeuTyr 151  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 256 AATGTTCCAATTATGAGAGAAAGACT----- 230  
QY 152 ProArgAspGlyValLeuLysGlyLulleHisLysAlaLeuLysLeuLysAspGlyGly 171  
DB -----CTGTGCTTGAAGGAATAAG 209  
QY 172 HisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysProValGlnLeuProGly 191  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 208 CATCACCGATGTAACCTTCAGAAAGTACTTACAAAGCCCAAGAGATGTGTTGCCAGAA 149  
QY 192 TyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleVal 211  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 148 TATCACTTTGTTGACCAACCGAATTGAGATATTAAAGTCATGACAAAGATTACAAACGTG 89  
QY 212 GluGlnTyrGluArgThrGlnGlyArg 220

DB 88 GTGCTCTATGAGATGCTGCTGCCGC 62

RESULT 2  
DR987866 556 bp mRNA linear EST 03-AUG-2005  
LOCUS JGI\_AOSF1131.fwd AOSF Montastraea faveolata adult colony  
DEFINITION Montastraea faveolata cDNA clone AOSF1131 5', mRNA sequence.  
ACCESSION DR987866 GI:71782572  
VERSION DR987866  
KEYWORDS EST.  
SOURCE Montastraea faveolata  
ORGANISM Montastraea faveolata  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;  
Favina; Faviidae; Montastraea.  
1 (bases 1 to 556)  
Schwarz,J.A., Brokstein,P., Manohar,C., Coffroth,M.A., Szmant,A.  
and Medina,M.  
Coral-Symbiodinium EST Project  
Unpublished (2005)  
Other ESTs: JGI\_AOSF1131.rev  
Contact: Schwarz, JA, Medina, M.  
Evolutionary Genomics  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925-296-5823  
Email: jaschwarz@lbl.gov  
cDNA library Preparation: DOE Joint Genome Institute:  
http://www.jgi.doe.gov

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished (2005)  
Other ESTs: JGI\_AOSF1131.rev  
Contact: Schwarz, JA, Medina, M.  
Evolutionary Genomics  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925-296-5823  
Email: jaschwarz@lbl.gov  
cDNA library Preparation: DOE Joint Genome Institute:  
http://www.jgi.doe.gov

DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov  
Naming Conventions: EST name is generated by the concatenation of the JGI Clone Id and the direction of sequencing. The suffix '.fwd' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.  
Small Insert: Based upon one or more sequencing reads of this clone where vector sequence was present at both ends, this clone has been determined to contain a cDNA insert on the order of 600-1000 bases.  
Plate: AOSF 0009 row: f column: 20  
High quality sequence stop: 557.

FEATURES  
Source  
Location/Qualifiers

1.556  
/organism="Montastraea faveolata"  
/mol\_type="mRNA"  
/db\_xref="taxon:48498"  
/clone="AOSF1131"  
/dev\_stage="Adult colony"  
/lab\_host="Electromax DH10B"  
/clone\_lib="AOSF Montastraea faveolata adult colony"  
/note="Vector: pDNR-LIB; Site 1: SfiI; Site 2: SfiI; The library was prepared from total RNA using the Creator SMART cDNA Library Construction Kit with the LD-PCR method to amplify the cDNA. Amplified cDNA was digested with SfiI, size selected for >400bp, and ligated into the pDNR-LIB vector. WARNING: this library contains a small percentage of cDNAs derived from the coral's symbiont, Symbiodinium sp."

ORIGIN

Alignment Scores:  
Pred. No.: 3.93e-45 Length: 556  
Score: 461.00 Matches: 83  
Percent Similarity: 62.96% Conservative: 36  
Best Local Similarity: 43.92% Mismatches: 46  
Query Match: 37.97% Indels: 24  
DB: 8 Gaps: 1

US-10-006-922A-12 (1-225) x DR987866 (1-556)

QY 32 GUGUGUGUGUAGArgProTyrGUGUGUHisAAnthrValLysLeuLysValThrLysGly 51  
DB ::::::::::::::::::::: ::::::::::: ::::: :::::  
DB 1 GACGGAGAGGCAAGCCTTTCGAGGAGAAACACAGATATAACCTTCAAGTCCAGAAAGGT 60  
QY 52 GlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysVal 71







Db 599 AAGAAGCAGCCGGTCTTCGTGATCCGCAAGTGCTACGTGAACGCTTCCAAGACCGAAATC 658  
QY 202 Thr 202  
Db 659 ACC 661  
RESULT 4  
BW802012 508 bp mRNA linear EST 23-MAY-2005  
LOCUS  
DEFINITION BW802012 Amphioxus Branchiostoma floridae unpublished cDNA library,  
larva whole animal Branchiostoma floridae cDNA clone bblv023d19 5',  
mRNA sequence.  
ACCESSION BW802012 GI:66410228  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM Branchiostoma floridae (Florida lancelet)  
Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.  
REFERENCE 1 (bases 1 to 508)  
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.  
TITLE Expressed genes in Branchiostoma floridae  
JOURNAL Unpublished (2005)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
LOCATION/Qualifiers  
1. .508  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="bblv023d19"  
/tissue\_type="whole animal"  
/dev\_stage="larva"  
/clone\_lib="Amphioxus Branchiostoma floridae unpublished  
cDNA library, larva whole animal"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.98e-16 Length: 508  
Score: 218.50 Matches: 52  
Percent Similarity: 51.41% Conservative: 21  
Best Local Similarity: 36.62% Mismatches: 50  
Query Match: 18.00% Indels: 19  
DB: 5 Gaps: 4  
US-10-006-922A-12 (1-225) x BW802012 (1-508)  
QY 18 MetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgPro 37  
Db 103 ATGTTGGGCTCCATCAACGGCCACGAGTTGCACTGGTGGCGGTGAAAGGCCACCCG 162  
QY 38 TyrGluGlyHisAsnThrValIleuLeuValThrIleGlyGlyProLeuProPheAla 57  
Db 163 AATGACGGGACGCTGAGACCAAGGTGCGGTCCACTTAAGGAGCC--CTGCCCTTCTCC 219  
QY 58 TrpAspIleLeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAla 77  
Db 220 CCCGTGATCCTGGCCCTTAACCTGGGGTACGGTACCAACGATCCTGCCCTTCCGGCC 279  
QY 78 AspIleProAspTyrIleuLeuSerPheProGluGlyPheIleuTrpGluArgValMet 97  
Db 280 GGGACCTCACCGTACGACGAGCCCATCACCAACGAGGGTACCAAAAGCACCGCAGTTC 339  
QY 98 AsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCys 117  
Db 340 AAGTTCGAGGACGGCGCGGTATGACCATCAAC----- 372  
/

QY 118 PheIleTyr-----LysValIle-----PheIleGlyValAsn 128  
Db 373 TTCGCGTACACCTACTCAGGAAACAAGATCAAGGAGAGATTCCACGCTGTGGATCCGGG 432  
QY 129 PheProSerAspGlyProValMetGlnIleuIleThrMetGlyTyrGluAlaSerThrGlu 148  
Db 433 TTCCCTGATGACGGCGCTGTGATGACCAACTCACTCCAGCAGCAGATCATCAAGTGGAG 492  
QY 149 ArgLeu 150  
Db 493 AGGCTG 498  
RESULT 5  
BW736869 646 bp mRNA linear EST 09-AUG-2005  
LOCUS  
DEFINITION BW736869 Amphioxus Branchiostoma floridae unpublished cDNA library,  
egg whole animal Branchiostoma floridae cDNA clone bbeg003b16 5',  
mRNA sequence.  
ACCESSION BW736869 GI:66323499  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM Branchiostoma floridae (Florida lancelet)  
Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.  
REFERENCE 1 (bases 1 to 646)  
AUTHORS Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.  
TITLE Expressed genes in Branchiostoma floridae  
JOURNAL Unpublished (2005)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
If you want to have a cDNA clone for this EST or if you have any  
questions, please send an e-mail to Nori Satoh  
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou  
(yutaka@ascidian.zool.kyoto-u.ac.jp).  
LOCATION/Qualifiers  
1. .646  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="bbeg003b16"  
/tissue\_type="whole animal"  
/dev\_stage="egg"  
/clone\_lib="Amphioxus Branchiostoma floridae unpublished  
cDNA library, egg whole animal"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.26e-15 Length: 646  
Score: 218.50 Matches: 54  
Percent Similarity: 48.28% Conservative: 30  
Best Local Similarity: 31.03% Mismatches: 85  
Query Match: 18.00% Indels: 5  
DB: 5 Gaps: 4  
US-10-006-922A-12 (1-225) x BW736869 (1-646)  
QY 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly 35  
Db 61 ATCCACCTTACCGGCTCCATCAACGGCCACGAGTTGCAATGTGGGGGAGAAAGGC 120  
QY 36 ArgProTyrGluGlyHisAsnThrValIleuLeuValThrIleGlyGlyProLeuPro 55  
Db 121 GACCGAAGCGCGCTCGGTGTGACCAAGCAAGAAATCCACCAAG--GGTCCCTGAAG 177  
QY 56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHis 75  
Db 178 TTCTCTCCCACTTATGATCCCCCACCCTGGGTACGGGTACTACCAAGTACCTCCCTAC 237  
/



[illegible]

RESULT 6	
BW736108	
LOCUS	BW736108
DEFINITION	BW736108 Amphioxus Branchiostoma floridae unpublished cDNA library, 639 bp mRNA linear EST 09-AUG-2005

ACCESSION	BW736108
VERSION	BW736108.1
KEYWORDS	GI:66322738
SOURCE	EST.
ORGANISM	Branchiostoma floridae (Florida lancelet) Branchiostoma floridae

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

REFERENCE	1 (bases 1 to 639)
AUTHORS	Yu, J., Holland, L. Z., Shin-i, T., Kohara, Y., Satou, Y. and Satch, N.
TITLE	Expressed genes in <i>Branchiostoma floridae</i>
JOURNAL	Unpublished (2005)
COMMENT	Contact: Tadasu Shin-i

Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: [tshini@genes.nig.ac.jp](mailto:tshini@genes.nig.ac.jp)  
If you want to have a cDNA clone for this EST or if you have any  
questions, please send an e-mail to Nori Satoh  
([satoh@ascidian.zool.kyoto-u.ac.jp](mailto:satoh@ascidian.zool.kyoto-u.ac.jp)) and its cc to Yutaka Satou  
([yutaka@ascidian.zool.kyoto-u.ac.jp](mailto:yutaka@ascidian.zool.kyoto-u.ac.jp)).

FEATURES	Location/Qualifiers
Source	1. .639

```

/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="bgeg023022"
/tissue_type="whole animal"
/dev_stage="egg"
/clone_lib="Amphioxus Branchiostoma floridae unpublished
cdna library, egg whole animal"

```

**ORIGIN**

Alignment Scores:	1.64e-15	Length:	639
Pred. No.:	217.50	Matches:	54
Score:	48.85%	Conservative:	31
Percent Similarity:	31.03%	Mismatches:	84
Best Local Similarity:	17.92%	Indels:	5
Query Match:	5	Gaps:	4
DB:			

US-10-006-922A-12 (1-225) x BW736108 (1-639)

Qy	16	ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly	35
		:::     ::::     ::::     ::::     ::::     ::::	
Db	67	ATCCACCTCCACGGCTCCATCAACGGCCACGAGTTCCAGCTGTGGCGGAGGAAGGC	126
Qy	36	ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro	55
Db	127	GACCCGAGCGCCGGCTCGTGGTGACACACAGCAATCCACCAAG--GGTCCCTGAAG	183
Qy	56	PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis	75
		:::     ::::                             :::	
Db	184	TTCTCTCCCTACTTGATGATCCCCACCCTCGGGTACGGGTACTACCAAGTACCTCCCTAC	243
Qy	76	ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp	93
		:::     :::	
Db	244	CCG--GACGGACCCCTCGCTTCCAGGCCACCACATGTTGGAAGGAACAGGGTATGCAGTC	300
Qy	94	GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu	113
		:::	
Db	301	CACCGCGTGTGCACTTTGAGACGGAGGCAAGCTGTCCATCGAGTTTAAGTACTCCTAC	360
Qy	114	GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnProSerAspGly	133
		:::     :::	
Db	361	GAGGGTCCCATATCAAGGCCGACATGAAGTTCACGGGACCGGTTTCCCTGAGACGGG	420
Qy	134	ProValMetGlnLysLysThrMetGlyTyrGluAlaSer--ThrGluArgLeuTyrPro	152
		:::     :::     :::     :::     :::	
Db	421	CCGGTCATGACCAAGCCAGATTATGACCAAGACGGGTGCGTGTCCAGAAACCTTACCTT	480
Qy	153	ArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHis	172
		:::     :::	
Db	481	AACGACAACACCATCGTGACAGCTTGACCTGACCTACAACCTGCGAATGGGAAGCGC	540
Qy	173	TyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro	186
		:::	
Db	541	TACAAGGCCCAAGTAGCAGGCCACTATCATCTTCGGCAAGCCC	582

RESULT	7
BW738638	
LOCUS	BW738638
DEFINITION	BW738638 Amphioxus Branchiostoma floridae unpublished cDNA library,
	604 bp mRNA linear EST 09-AUG-2005

ACCESSION	BW738638	
VERSION	BW738638.1	GI:66325268
KEYWORDS	EST.	
SOURCE	Branchiostoma floridae	(Florida lancelet)
ORGANISM	Branchiostoma floridae	

REFERENCE	1 (bases 1 to 604)
AUTHORS	Yu, J., Holland, L. Z., Shin-i, T., Kohara, Y., Satou, Y. and Satoh, N.
TITLE	Expressed genes in <i>Branchiostoma floridae</i>
JOURNAL	Unpublished (2005)
COMMENT	Contact: Tadasu Shin-i

Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: [tsuhini@genes.nig.ac.jp](mailto:tsuhini@genes.nig.ac.jp)  
If you want to have a cDNA clone for this EST or if you have any  
questions, please send an e-mail to Nori Satoh  
([satoh@ascidian.zool.kyoto-u.ac.jp](mailto:satoh@ascidian.zool.kyoto-u.ac.jp)) and its cc to Yutaka Satou  
([yutaka@ascidian.zool.kyoto-u.ac.jp](mailto:yutaka@ascidian.zool.kyoto-u.ac.jp)).

## FEATURES

```
source
1. 604
/organism="Branchiostoma floridae"
/mol_type="rRNA"
/db_xref="taxon:7739"
/closure="bseg008a05"
/tissue_type="whole animal"
```



/dev\_stage="egg"  
/clone\_lib="Amphioxus Branchiostoma floridae unpublished  
cDNA library, egg whole animal"

ORIGIN

Alignment Scores:

Pred. No.:	2.01e-15	Length:	604
Score:	216.50	Matches:	54
Percent Similarity:	48.85%	Conservative:	31
Best Local Similarity:	31.03%	Mismatches:	84
Query Match:	17.83%	Indels:	5
DB:	5	Gaps:	4

US-10-006-922A-12 (1-225) x BW738638 (1-604)

QY	16	ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGly	35
DB	67	ATCCACCTCCACGGCTCCATCAACGGCCACGAGTTGACCTGGTGGCGGAGAAAGGC	126
QY	36	ArgProTyrGluGlyHisAsnThrValIleValIleValThrIleValThrIleValPro	55
DB	127	GACCCGAACGGCGCTCGCTGTGACCAACAGCGAAATCCACCAAG--GGTCCCTGAAG	183
QY	56	PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHis	75
DB	184	TTCTCTCCCTACTGATGATGCCCACTCGGGTACGGTACTACCACTACCTCCCTTAC	243
QY	76	ProAlaAspIleProAspTyrIleValIleValLeuSerPheProGlu-----GlyPheIle	93
DB	244	CCG--GACGGACCCCTCGCTTTCAGGCCACCATGTTGGAAGATCAGGGTATGCAGTC	300
QY	94	GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu	113
DB	301	CATCGCGTGTTCGACTTTGAGAAGCGAGGCAAGCTGTCATCGAGTTTAAGTACTCTTAC	360
QY	114	GlnAspGlyCysPheIleTyrIleValIleValPheIleGlyValAsnPheProSerAspGly	133
DB	361	GAGGGTTCATATATCAAGCGGACATGAAGTTCAAGGAAACGGGTTCCCTGAGACGGG	420
QY	134	ProValMetGlnIleValThrMetGlyTyrGluAlaSer---ThrGluArgLeuTyrPro	152
DB	421	CCGGTCATGACAAAGCCAGATTGTGCAGCAAGACGGCTGTCCAAAGAACCTTACCTT	480
QY	153	ArgAspGlyValLeuIleValIleHisIleValIleValIleValIleValIleValIle	172
DB	481	AACGACAACACCATCGTGACAGACCTTGACTGACCTTACCACTGCAAGATGGGAAGCCG	540
QY	173	TyrLeuValGluPheIleTyrIleTyrMetAlaIleValPro	186
DB	541	TACAAAGGCCAAAGTGACGACCACTACATCTTCGGCAAGCCC	582

RESULT 8

BW741633 606 bp mRNA linear EST 09-AUG-2005  
LOCUS BW741633 Amphioxus Branchiostoma floridae unpublished cDNA library,  
DEFINITION egg whole animal Branchiostoma floridae cDNA clone bbeg040g05 5',  
mRNA sequence.

ACCESSION BW741633 GI:66328281  
VERSION BW741633.1  
KEYWORDS EST.  
SOURCE Branchiostoma floridae (Florida lancelet)  
ORGANISM Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.

REFERENCE 1 (bases 1 to 606)  
AUTHORS Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.  
TITLE Expressed genes in Branchiostoma floridae  
JOURNAL Unpublished (2005)  
COMMENT Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856

Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp  
If you want to have a cDNA clone for this EST or if you have any  
questions, please send an e-mail to Nori Satoh  
(satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou  
(yutaka@ascidian.zool.kyoto-u.ac.jp).

FEATURES

SOURCE

1. .606  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="bbeg040g05"  
/issue\_type="whole animal"  
/dev\_stage="egg"  
/clone\_lib="Amphioxus Branchiostoma floridae unpublished  
cDNA library, egg whole animal"

ORIGIN

Alignment Scores:

Pred. No.:	2.01e-15	Length:	606
Score:	216.50	Matches:	54
Percent Similarity:	48.28%	Conservative:	30
Best Local Similarity:	31.03%	Mismatches:	85
Query Match:	17.83%	Indels:	5
DB:	5	Gaps:	4

US-10-006-922A-12 (1-225) x BW741633 (1-606)

QY	16	ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGly	35
DB	69	ATCCACCTTCACGGCTCCATCAACGGCCACGAGTTGCACATGTTGGGGGAGAAAGGC	128
QY	36	ArgProTyrGluGlyHisAsnThrValIleValIleValThrIleValThrIleValPro	55
DB	129	GACCCGAACGGCGCTCGCTGTGACCAACAGCGAAATCCACCAAG--GGTCCCTGAAG	185
QY	56	PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHis	75
DB	186	TTCTCTCTTACTGATGATCCCCACCTCGGGTACGGGTACTACCACTACCTCCCTTAC	245
QY	76	ProAlaAspIleProAspTyrIleValIleValLeuSerPheProGlu-----GlyPheIle	93
DB	246	CCG--GACGGACCCCTCGCTTTCAGGCCACCATGTTGGAAGATCGGGTATTCAAGTC	302
QY	94	GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu	113
DB	303	CACCGGTGTTGATTTGAAAGCGAGGCAAGCTGCCATCGAGTTTAAGTACTCCTTAC	362
QY	114	GlnAspGlyCysPheIleTyrIleValIleValPheIleGlyValAsnPheProSerAspGly	133
DB	363	GAGGGTTCATATTAAGCGGACATGAAGTTCAAGGAAACGGGTTCCCTGAGACGGC	422
QY	134	ProValMetGlnIleValThrMetGlyTyrGluAlaSer---ThrGluArgLeuTyrPro	152
DB	423	CCGGTCATGACCAAGCCAGATTGTGCAGCAAGACGGCTGCTCCAAAGAACCTTACCTC	482
QY	153	ArgAspGlyValLeuIleValIleHisIleValIleValIleValIleValIleValIle	172
DB	483	AACGACAACACCATCGTGACAACTTGACTGACCTTACAATCTGCAGAACGAAAGCCG	542
QY	173	TyrLeuValGluPheIleTyrIleTyrMetAlaIleValPro	186
DB	543	TACAGAGCCGAGTGACGACCACTACATCTTCGACAAAGCCC	584

RESULT 9

BI379468 637 bp mRNA linear EST 26-AUG-2003  
LOCUS BI379468  
DEFINITION BFLG1 001099 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG  
or MPMGP498) Branchiostoma floridae cDNA clone MPMGP498O1918 5',  
mRNA sequence.

ACCESSION BI379468  
VERSION BI379468.1 GI:30914617  
KEYWORDS EST.



SOURCE Branchiostoma floridae (Florida lancelet)  
ORGANISM Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.  
REFERENCE 1 (bases 1 to 637)  
AUTHORS Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.  
TITLE New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes  
JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)  
PUBMED 12799346  
COMMENT Contact: Panopoulou G  
laboratory 145, dept.Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inmestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTTATGCTTCGGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGAAGGGGATGTG 3' (M13FSP)  
Insert Length: 1200 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGAATTCCCGCGT-3' pSport3/86  
High quality sequence stop: 637.  
location/Qualifiers  
1. 637  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="MPMGp49801918"  
/issue\_type="whole embryo"  
/dev\_stage="5-6 hrs (gastrula stage)"  
/lab\_host="E.coli, XL1 blue"  
/clone\_lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLg or MPMGp498)"  
/note="Vector: pSport1; Site\_1: SalI, KpnI, EcoRI (5'); Site\_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTCTAGATCCGCGAGCGCGCC (T)15-3' and a SalI 5'-TCGACCCACGCGCTCCG-3'adapters (Gibco BRL)."  
ORIGIN  
.  
Alignment Scores:  
Pred. No.: 2,16e-15 Length: 637  
Score: 216.50 Matches: 53  
Percent Similarity: 48.57% Conservative: 32  
Best Local Similarity: 30.29% Mismatches: 85  
Query Match: 17.83% Indels: 5  
DB: 3 Gaps: 4  
US-10-006-922A-12 (1-225) x B1379468 (1-637)  
QY 15 LysValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlu 34  
Db 79 GAGATCCACCTCCACGGTTCATCAACGCCACGAGTTGACCTGCTGTGTGAAAA 138  
QY 35 GlyArgProTyrGluGlyHisAsnThrValIleuLysValThrLysGlyGlyProLeu 54  
Db 139 GCGGACCCAAACGCCGGCTCGCTGTGACCAACAGCAATCCACCCAG---GGTCCCTG 195

QY 55 ProPheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLys 74  
Db 196 AAGTTCCTCCCCACCTTGATGATCCCCACCTCGGGTACGGGTACTACCAAGTACTCCCC 255  
QY 75 HisProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLys 92  
Db 256 TACCG--GACGAGCCCTCGCCCTTCCAGGCCCACTATGTGAAGATCGGGGTATACA 312  
QY 93 TrpGluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSer 112  
Db 313 GTCCACCGCGTGTGTTGACTTCGAAGACGAGGCAAGCTGTCCATCGAGTTCAGTACTCC 372  
QY 113 LeuGlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSeraAsp 132  
Db 373 TACGAGGTTCCCATATTAAAGCCGACATGAAGTTCAAGGAACGGTTCCCTGAGGAT 432  
QY 133 GlyProValMetGlnLysLysThrMetGlyTyrGluAlaSer---ThrGluArgLeuTyr 151  
Db 433 GGGCCGGTCATGACCAAGCCAGATTGTGCAGCAGAGCGGCTGCTGCCAAGAACAACCTAC 492  
QY 152 ProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGly 171  
Db 493 CTCACGACACACACCATCTGGACCAACTTCGACTGCACTTACAACCTGCAGATGGGAAG 552  
QY 172 HisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
Db 553 CGCTACAGGGCCCGAGTGACGAGCCCACTACATCTTCGACAAAGCCC 597  
RESULT 10  
B1387917 687 bp mRNA linear EST 26-AUG-2003  
LOCUS BFL26\_002620 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531) Branchiostoma floridae cDNA clone MPMGp531N0767 5', mRNA sequence.  
ACCESSION B1387917  
VERSION B1387917.1 GI:30922753  
KEYWORDS EST.  
SOURCE Branchiostoma floridae (Florida lancelet)  
ORGANISM Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.  
REFERENCE 1 (bases 1 to 687)  
AUTHORS Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.  
TITLE New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes  
JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)  
PUBMED 12799346  
COMMENT Contact: Panopoulou G  
laboratory 145, dept.Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inmestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTACACTTTATGCTTCGGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGAAGGGGATGTG 3' (M13FSP)  
Insert Length: 1 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGAATTCCCGCGT-3' pSport3/86



High quality sequence stop: 687.  
 Location/Qualifiers  
 1..687  
 /organism="Branchiostoma floridae"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7739"  
 /clone="MPMG531N0767"  
 /tissue\_type="whole embryo"  
 /dev\_stage="26 hrs (neurula stage)"  
 /lab\_host="E. coli, XLI blue"  
 /clone\_id="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMG531)"  
 /note="Vector: pSport1 (Gibco BRL); Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-PGACTAGTTCTAGATCGGAGCGGCCGCC (T)15-3' and a SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

ORIGIN

Alignment Scores:  
 Pred. No.: 2.41e-15 Length: 687  
 Score: 216.50 Matches: 54  
 Percent Similarity: 48.85% Conservative: 31  
 Best Local Similarity: 31.03% Mismatches: 84  
 Query Match: 17.83% Indels: 5  
 DB: 3 Gaps: 4

US-10-006-922A-12 (1-225) x BI387917 (1-687)

```

QY      16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly 35
       :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
DB      99 ATCCACCTCCACGGCTCCATCAACGCCACGAGTTCGACATGGTGGCGGAGGAAAGGC 158
QY      36 ArgProTyxGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeu 55
       |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB      159 GACCCGAAAGCCGGCTCGCTGCTGTGACCAAGCAAGCAATCCACCAAG--GATCCCTGAAG 215
QY      56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyxGlySerLysValTyxValLysHis 75
       ||::: ::::: ::::: |||| |||| |||| |||| |||| ||||
DB      216 TTCTCTCCCTACTTGATGATCCCCACCTCGGGTACGGTACTACAGTACCTCCCTAC 275
QY      76 ProAlaAspIleProAspTyxLysLysLeuSerPheProGlu-----GlyPheLysTrp 93
       |||| |||| |||| :::: :::: :::: |||| |||| ||||
DB      276 CCG--GACGGACCCCTCGCCTTCCAGGCCACCATGTTGGAAGAACAGGGTATGCAGTC 332
QY      94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113
       ||||| ||||| ::::| ||||| ||||| ::::| |||||
DB      333 CATCGCGTGTTCGACTTTGAAGAAGGAGGCAAGCTGTCATCGAGTTTAAGTACTCCTAC 392
QY      114 GlnAspGlyCywPheIleTyxLysValLysPheIleGlyValAsnPheProSerAspGly 133
       ::: :::: ::::: ||||| ||||| ||||| ||||| |||||
DB      393 GAGGGTTCCTCATATCAAGGCGGACATGAAGTTCAGGGAACCGGTTTCCCTGAGGACGGG 452
QY      134 ProValMetGlnLysValThrMetGlyTrpGluAlaSer--ThrGluArgLeuTyxPro 152
       ||||| ||||| :::: :::: ::::: ||||| ||||| |||||
DB      453 CCGGTCAATGACCAAGCAAGTGTGCAACGAGGACGGCTGCTGTCGAAGAACAACCTACCTT 512
QY      153 ArgAspGlyValLeuLysGlyGluIleHisLysValAlaLeuLysValAspGlyGlyHis 172
       |||| :::: ::::: ||||| ||||| ||||| ||||| |||||
DB      513 AACGACAAACACCATCGTGACAGCTTCGACTGCACTTAAACCTGCAATGGGAAGCGC 572
QY      173 TyrLeuValGluPheLysSerIleTyxMetAlaLysLysPro 186
       |||| :::: ::::: ||||| ||||| ||||| |||||
DB      573 TACAAGGCCCAAGTACGAGCCACTATCTTCCGCAAGCCC 614
  
```

RESULT 11

LOCUS BI386699 643 bp mRNA linear EST 26-AUG-2000

DEFINITION BFL26\_001402 Amphioxus 26hr cDNA library (Name convention: BFL26 o MPMG531) Branchiostoma floridae cDNA clone MPMG531L1370 5', mRNA sequence.

ACCESSION BI386699 BI386699

VERSION BI386699.1 GI:30921604

KEYWORDS EST.  
SOURCE Branchiostoma floridae (Florida lancelet)  
ORGANISM Branchiostoma floridae  
AUTHORS Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
REFERENCE Branchiostoma.  
TITLE 1 (bases 1 to 643)  
JOURNAL Panopoulou,G., Hennig,S., Groth,D., Krause,A., Pousetka,A.J.,  
PUBMED Herwig,R., Vingron,M. and Lehrach,H.  
COMMENT New evidence for genome-wide duplications at the origin of  
vertebrates using an amphioxus gene set and completed animal  
genomes  
Genome Res. 13 (6A), 1056-1066 (2003)  
12799346  
Contact: Panopoulou G  
Laboratory 145, dept.Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONFP) to reduce sequencing redundancy. According to the ONFP  
procedure, clones giving the same hybridisation pattern with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well the coordinates of the  
rest of the clones assigned to the same fingerprint cluster as the  
clone from which the above EST is generated is available at the  
amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary  
Database of the German Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGGCTTTACACTTTATGCTTCGCGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)  
Insert length: 1 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGGAATCCCGGGT-3' pSPORT3/86  
High quality sequence stop: 643.  
location/Qualifiers  
1..643  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="MPMGp531L1370"  
/tissue\_type="whole embryo"  
/dev\_stage="26 hrs (neurula stage)"  
/lab\_host="E.coli, XL1 blue"  
/clone\_lib="Amphioxus 26hr cDNA library (Name convention:  
BFL26 or MPMGp531)"  
/note="Vector: pSPORT1 (Gibco BRL); Site\_1: SalI, XpnI,  
EcoRI (5'); Site\_2: NotI, BamHI, HindIII (3'); OligodT  
primed and directionally cloned in pSPORT1 vector using a  
NotI (5'-pGACTAGTCTAGATCGGAGCGCGCCC (T)15-3' and a  
SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.69e-15 Length: 643  
Score: 212.50 Matches: 56  
Percent Similarity: 47.87% Conservative: 34  
Best Local Similarity: 29.79% Mismatches: 87  
Query Match: 17.50% Indels: 11  
DB: 3 Gaps: 5  
US-10-006-922A-12 (1-225) x BI386699 (1-643)  
Oy ValaArgMetGluGlyThrValaAenGlyHisGluPheGluIleGluGlyGluGly 35  
Db 81 ATCACTTCAACGGCTCATCAACGGCCACGAGTTGCACATGATGGGGGAGAAAGGC 140  
Oy 36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro 55  
Db 141 GACCCGAACGCCGGCTCGCTGGTGAACCAACGGAATCCACCAAG---GGTCCCTGAAG 197



Oy		56	PheAlATrAspRIleLeuSerProGlnPheGlnTyrGIyserylValTyrValIysHis	75
Db		198	TTCTCTCCCTACTGTGATGATCCCCACCCTCGGGTAAGGGTACTACCAAGTACCTCCCCTAC	257
Oy		76	ProAlAspRIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp	93
Db		258	CCG--GACGGAACCTCGCCTTTCAGGCCTCATGTTGGAAGGATCGGGATATGACATC	314
Oy		94	GluArgValMetAsnPhelGluAspGlyGlyValValThrValThrglnAspSerSerIeu	113
Db		315	TACCGCGTGTTGCACTTTGAAGAAGCGAGGCAAGCTGACTAACCGAGTTTAAGTACTCTTAC	374
Oy		114	GlnAspGlyCySPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly	133
Db		375	GAGGATTCCCAATCAAGCCGACATGAAAGCTGATGGGAAGCGTTTCCCTGACGACGGC	434
Oy		134	ProValMetGlnLysLysThrMetGlyTrpGluAlaSer---ThrgluArgLeuTyrPro	152
Db		435	CCAGTCATGACCAAGCCAGATTTGTGCACCAGACGGCGTGCATGCCAAGAAGACGTATCTT	494
Oy		153	ArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHis	172
Db		495	AACAACAACAACCATCGTGGACAGCTTCGACTGGAGTTACAACCTGCAGAATGGAAAGCGC	554
Oy		173	TyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro-----	186
Db		555	TACAGGGCCCGAGTGTCAAGCCACTACATCTTCGACAAAGCCCTTTCAGCCGATCTCATG	614
Oy		187	ValGlnLeuProGlyTyrTyrTyr	194
Db		615	ATAAAGCAGCCCGGTCTTCGTGTAC	638

RESULT 12	
LOCUS	675 bp mRNA linear EST 26-AUG-2003
BI382638	
DEFINITION	BFLG2 000815 Amphioxus 5-6 hrs CDNA library (Name convention: BFLG or MPMGp498) Branchiostoma floridae CDNA clone MPMGp498F0258 5', mRNA sequence.
ACCESSION	BI382638
VERSION	BI382638.1 GI:30917704

SOURCE ORGANISM	BRANCHIOSTOMA
Branchiostoma floridae (Florida lancelet)	Branchiostoma floridae
Branchiostoma floridae	Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.	Branchiostoma.

REFERENCE	1 (bases 1 to 675)
AUTHORS	Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.
TITLE	New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes
JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
PUBMED	12799346
COMMENT	Contact: Panopoulou G

JOURNAL  
PUBMED  
COMMENT

Genome Res. 13 (6A), 1056-1066 (2003)  
12799346  
Contact: Panopoulou G  
laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoulou@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphiToxus project site at: <http://www.molgen.mpg.de/amphiToxus/> clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR primers  
FORWARD: 5' CCCGAGCTTTACACTTATGCTTCGGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGAGTGTG 3' (M13FSP)  
Insert Length: 1 Std Error: 0.00  
Seq primer: 5'-CCGTCGGAGTATCCCGGGT-3' pSport3/86  
High quality sequence stop: 675.

**FEATURES**  
**Source**

```

1. .675
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MPMGp498F0258"
/tissue_type="whole embryo"
/dev_stage="5-6 hrs (gastrula stage)"
/lab_host="E.coli, XL1 blue"
/clone_1b="Amphioxus 5-6 hrs cDNA library (Name
convention: BFLG or MPMGP498)"
/notes="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5');
Site 2: NotI, BamHI, HindIII (3'); OligodT primed and
directionally cloned in pSport1 vector using a NotI
(5'-GCACTAGTCTTAGATCGGAGCGCGCGCC (T)15-3' and a SalI 5'-
TCGACCCACGCGCTCCG-3' adapters (Gibco BRL)."
```

**ORIGIN**

Alignment Scores:	
Pred. No.:	7.17e-15
Score:	212.50
Percent Similarity:	49.43%
Best Local Similarity:	31.03%
Query Match:	17.50%
DB:	3
Length:	675
Matches:	54
Conservative:	32
Mismatches:	83
Indels:	5
Gaps:	4

US-10-006-922A-12 (1-225) x BI382638 (1-675)

QY 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly 35  
 Db 110 ATCCACCTTCACGGCTCCATCAACGCCACGAGTTCCACATGTTGGGGGAGAAAGGC 169  
 QY 36 ArgProTyrGluGlyHisAsnThrValLysLeuValThrLysGlyGlyProLeuPro 55  
 Db 170 GACCCGAAACGCCGGCTCGCTGGTGCACACAGCGAAATCCACCAAG--GGTCCCTGAAG 226  
 QY 56 PheAlaTyrAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75  
 Db 227 TTCTCCCTACTTGAATGATCCCCCACTCGGGTACGGGTACTACCAAGTACCTCCCTAC 286  
 QY 76 ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu----GlyPheLysTrp 93  
 Db 287 CCG---GACGAGACCCCTCGCCTTTCAGGCTCCATGTTGGAAGATCGGGGTATGCAGTC 343  
 QY 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113  
 Db 344 TACCCGCGTGTTCGACTTTCAGAGACGGAAGCAAGCTGACTACCGAGTTTAAGTACTCCTAC 403  
 QY 114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly 133  
 Db 404 GAGGGTCCCATATCAAGGCCGACATGAAGCTGATGGGAAGCGGTTTCCCTGACGACGGC 463  
 QY 134 ProValMetGlnLysLysThrMetGlyTyrGluAlaSer--ThrGluArgLeuTyrPro 152  
 Db 464 CCAGTCATGACCAAGCCAGATTGTGCACGAGACGGCTGCTGCCAAGAAGACGTATCTT 523  
 QY 153 ArgAspGlyValLeuLysGlyGluIleHisLysValAlaLeuLysLeuValAspGlyGlyHis 172  
 Db 524 AACACACACACACCATCGTGACAGCTTCGACTGGAGTTACAACTGCAGATGGGAAGCGC 583  
 QY 173 TyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
 Db 584 TACAGAGCCCGAGTGTCAAGCCCACTACATCTTCGACAAAGCCC 625

RESULT 13				
BW883955				
LOCUS				
BW883955	590 bp	mRNA	linear	EST 24-MAY-2005



DEFINITION	BW883955 Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal Branchiostoma floridae cDNA clone bne135j09 5', mRNA sequence.		
ACCESSION	BW883955		
VERSION	BW883955.1 GI:66498632		
KEYWORDS	EST.		
SOURCE	Branchiostoma floridae (Florida lancelet)		
ORGANISM	Branchiostoma floridae Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.		
REFERENCE	1 (bases 1 to 590)		
AUTHORS	Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.		
TITLE	Expressed genes in Branchiostoma floridae		
JOURNAL	Unpublished (2005)		
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. Location/Qualifiers		
FEATURES	1..590		
source	/organism="Branchiostoma floridae" /mol_type="mRNA" /db_xref="taxon:7739" /clone="bne135j09" /tissue_type="whole animal" /dev_stage="neurula" /clone_lib="Amphioxus Branchiostoma floridae unpublished cDNA library, neurula whole animal"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.04e-14	Length:	590
Score:	210.50	Matches:	54
Percent Similarity:	48.85%	Conservative:	31
Best Local Similarity:	31.03%	Mismatches:	84
Query Match:	17.34%	Indels:	5
DB:	5	Gaps:	4
US-10-006-922A-12 (1-225) x BW883955 (1-590)			
QY	16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGly	35	
DB	70 ATCCACCTTCACGGCTCCATCAACGCCACGAGTTCCACATGGTGGGGAGGAAAGGC	129	
QY	36 ArgProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuPro	55	
DB	130 GACCCGAACGCCGCTCGCTGGTGACACAGCGAAATCCACAAG--GGTCCCTGAAG	186	
QY	56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis	75	
DB	187 TTCTCTCCCTACTTGATGATCCCCACCTCGGGTACGGGTACTACCAAGTACCTCCCCCTAC	246	
QY	76 ProAlaAspIleProAspTyrLysLysLeuSerPheProGlu-----GlyPheLysTrp	93	
DB	247 CCG--GACGGACCCCTCGCCTTTCCAGGCCCTCCATGTTGGAAGATCGGGGTATGCAGTC	303	
QY	94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu	113	
DB	304 TACCGCGTGTTCGACTTTGAAGACGGAAGCAAGCTGACTACCGAGTTTAAGTACTCCTAC	363	
QY	114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly	133	
DB	364 GAGGGTTCACATATCAAGGCCACATGAAGCTGATGGGAAGCGGTTTCCCTGACGACGGC	423	
QY	134 ProValMetGlnLysLysThrMetGlyTyrGluAlaSer---ThrGluArgLeuTyrPro	152	
DB	424 CCAATCATGGCCAGCCAGATTGTGACCAAGACGGCTGCGTGTCCAAAGAACGATATCTT	483	
QY	153 ArgAspGlyValLeuLysGlyGluIleHisLysValAlaLeuLysLeuLysAspGlyGlyHis	172	

```

Db      484 AACACAACACCATCGTGGACAGCTTCGACTGGAGTTACAACCTGCAGAATGGGAAGCC 543
Qy      173 TyrLeuValGlupHelysSerIleTyMetAlalysPro 186
          ||| |::| |::| |::| |::|
Db      544 TACAGGGCCCCGAGTGTCGAGCAGTACATCTTCGACAAGCCC 585

RESULT 14
BM893733                                812 bp    mRNA       linear   EST 24-MAY-200
LOCUS  BM893733                          Amphioxus Branchiostoma floridae unpublished cDNA library
DEFINITION
neurula whole animal Branchiostoma floridae cDNA clone bnei23k07
5', mRNA sequence.
ACCESSION  BM893733
VERSION    BM893733.1  GI:66511577
KEYWORDS  EST.
SOURCE    Branchiostoma floridae (Florida lancelet)
ORGANISM  Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
Branchiostoma.
REFERENCE  1 (bases 1 to 812)
AUTHORS  Yu,J., Holland,L.Z., Shin-I,T., Kohara,Y., Satou,Y. and Satoh,N.
TITLE     Expressed genes in Branchiostoma floridae
JOURNAL   Unpublished (2005)
COMMENT   Contact: Tadasu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp.
FEATURES
         source
           1..812
             /organism="Branchiostoma floridae"
             /mol_type="mRNA"
             /db_xref="taxon:7739"
             /clone="bnei23k07"
             /tissue_type="whole animal"
             /dev_stage="neurula"
             /clone_lib="Amphioxus Branchiostoma floridae unpublished
             cDNA library, neurula whole animal"

ORIGIN

Alignment Scores:
Pred. No.:      1.63e-14      length:      812
Score:          210.50        Matches:      60
Percent Similarity: 46.92%    Conservative: 39
Best Local Similarity: 28.44% Mismatches:    91
Query Match:    17.34%       Indels:       21
DB:              5           Gaps:         7

US-10-006-922A-12 (1-225) x BM893733 (1-812)

Qy      16 ValArgMetGlUgLyThrValaAngLyHisGlupheglutleglUgLyglUgLy 35
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      60 ATCCACCTTCACGGCTCATCAACGCCCAAGATTGACATGTGGGGGAGGAAAAAGCC 119
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy      36 ArgProTyrgLugLyHisAnthrVallysleuYsValThrxlysglygYProleuPro 55
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      120 GACCCGAACGCCGCTGCTGCTGACCAACAGCGAAATCCACCAGGGTCC--CTGAAG 176
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy      56 pheAlatPraSpIlleleuSerProGlnPheglIntyrGlyserlysvAltyrValylsHis 75
          |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      177 TTCTCTCCCTACTTGATGATGCCCACTCGGGTAGCGGTACTACCAAGTACCTCCCTAC 236
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy      76 ProAlaaspIleProaspTyrlYlsBlysleuSerPheProglu-----GlyPhelysTrp 93
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      237 CCG--GACGAGCCCTCGCCTTTCAGACCTCCATGTTGGAAGGATCGGGGTATGCAGTC 293
          ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Qy      94 GluArgValmetAsnphegliuaspglyglYvalValThrxvalThrglnaspserserleu 113
          |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      294 TACCGCGTGTTCGACTTTGAAGACGGAAGCAAGCTGACTACCGAGTTTAAGTACTCCTAC 353
          |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      114 GluAspGlyCySPheIleTyrlYlsVallysPheIleGlyValaSnPheProSerAspGly 133

```



DB 354 GAGGGTCCCATATCAAGGCCGACATGAAGCTGATGGGAAGCGGTTTCCCTGACGAGCGC 413  
QY 134 ProValMetGlnLysLysThrMetGlyTrpGluAlaSer----ThrGluArgLeuTyrPro 152  
Db 414 CCAGTCATGACCAAGCCAGATGTTCAGACGAGCGGCTCGGTGTCACAGAAGACGTATCTT 473  
QY 153 ArgAspGlyValLeuLysGlyGluIleHisLysValAlaLeuLysLeuLysAspGly----- 170  
Db 474 AACGACAACACCATCGTGACAGCTTCGACTGAGTTACAACCTTGACAGAATGGGAAGCGC 533  
QY 171 -----GlyHisTyrLeuValGlu-----PheLysSerIleTyrMet 182  
Db 534 TACAGGGCCCGAGTGTGAGCCACTACATCTTCGACAAGCCCTTTTCAGCCGATCTCATG 593  
QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
Db 594 AAGAACGACCGCGTG-----TTCGTGTACCGGAAGTGCCACGTGAAG 635  
QY 203 SerHisAsnGluAspTyrThrIleValGluGln 213  
Db 636 GCTTCCAAGACCGAAGTCACCCTGACGAGAGG 668

RESULT 15  
AL044652 790 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZp434P092\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
DEFINITION DKFZp434P092 5', mRNA sequence.  
ACCESSION AL044652  
VERSION AL044652.1 GI:5432867  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 790)  
AUTHORS Ansoerge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and  
Wiemann,S.  
TITLE EST (Ansoerge, Benes, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: MIPS

MIPS  
ingolstaedter landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No si sequence available.  
This clone (DKFZp434P092) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1. .790  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp434P092"  
/issue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="434 (synonym: htes3)"  
/note="Vector: pSport1, Site\_1: NotI; Site\_2: SalI"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.07e-14 Length: 790  
Score: 209.50 Matches: 55  
Percent Similarity: 47.25% Conservative: 48  
Best Local Similarity: 25.23% Mismatches: 102  
Query Match: 17.26% Indels: 13

DB: 1 Gaps: 6  
US-10-006-922a-12 (1-225) x AL044652 (1-790)  
QY 16 ValArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGly 35  
Db 32 GTGGAACCTGGATGGCGGATGTGAATGGGACACAATTTTCTGTACGGGAGAGGGTGAAGT 91  
QY 36 ArgProTyrGluGlyHisAsnThrValLysLeuValThrLysGlyGlyProLeuPro 55  
Db 92 GATGCCACATACGGAAGCTCAACCTGAATTTCATCTGCACCACT--GGAAAGCTCCCT 148  
QY 56 PheAlaTrpAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHis 75  
Db 149 GTGCCATGGCCAACACTGTGCTACCTTCACTATGGCGTGCAGTGCTTTTCCAGATAC 208  
QY 76 ProAlaAspIlePro-----AspTyrLysLysLeuSerPheProGluGlyPheLysTrp 93  
Db 209 CCAGACCATATGAAGCAGCATGACTTTTCAAGAGCGCCATGCCAGGCTATGTGCAG 268  
QY 94 GluArgValMetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeu 113  
Db 269 GAGAGAACCATCTTTTTCAAAGATGACGGGAACCTACAAGACCCGCGCTGAAGTCAAGTTC 328  
QY 114 GlnAspGlyCysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGly 133  
Db 329 GAAGGTGACACCCCTGTGATATGAATCGAGCTGAAGGGCATTTGAAGAGGATGGA 388  
QY 134 ProValMetGlnLysLysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArg 153  
Db 389 AACATTCTCGGCCACACAAG--CTGGAATACAACTATACTCCCACAATGTGTACATCATG 445  
QY 154 AspGlyValLeuLysGlyGluIleHisLysValAlaLeuLysLeuLys-----Asp 169  
Db 446 GCCGACAACAAGAAATGGCATCAAGTCAACTTCAAGATCAGACACAACATTGAGGAT 505  
QY 170 GlyGlyHisTyrLeuValGluPhe-----LysSerIleTyrMetAlaLysLysProVal 187  
Db 506 GGATCCGTGACAGCTGGCCGACCATTAACAAGACACTCCAAATCGGCGAGCCCTGTG 565  
QY 188 GlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr-----SerHis 204  
Db 566 CTCCTCCAGACAACCATTAACCTGTCCACCCAGTCTGCCCTGTCTAAAGATCCCAACGAA 625  
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArgHisHis 222  
Db 626 AAGAGAGCACCATGTGCTGTGAGTTGTGACCGGCTGCTGGATCAACACAT 679

Search completed: January 12, 2006, 12:23:29  
Job time : 3750 secs



**THIS PAGE BLANK (USPIO)**



Gencore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 09:45:12 ; Search time 179 Seconds  
(without alignments)  
2234.365 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNV1KEFMRFKVRMEG.....EDYTIIVEQYERTGRRHLLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 segs, 888780828 residues  
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlj  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_10012006\_162409\_18947/app\_query.fasta\_1.391  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922\_@CGN\_1\_1\_290\_@runat\_10012006\_162409\_18947 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfile81.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1214	100.0	678	3	US-09-459-956-6	Sequence 6, Appli
2	1214	100.0	859	3	US-09-866-538-11	Sequence 11, Appl
3	1214	100.0	859	3	US-09-865-291-11	Sequence 11, Appl
4	1210	99.7	723	3	US-10-152-296-1	Sequence 1, Appli
5	1210	99.7	6984	3	US-10-001-189-45	Sequence 45, Appl
6	729.5	60.1	699	3	US-09-459-956-5	Sequence 5, Appli
7	681	56.1	860	3	US-10-244-779-1	Sequence 1, Appli
8	667	54.9	801	3	US-09-459-956-7	Sequence 7, Appli
9	495	40.8	690	3	US-09-459-956-2	Sequence 2, Appli

10	486.5	40.1	1079	3	US-09-609-161B-15	Sequence 15, Appl
11	486.5	40.1	1079	3	US-09-626-581D-64	Sequence 64, Appl
12	486.5	40.1	1079	3	US-09-415-765B-64	Sequence 64, Appl
13	486.5	40.1	1079	3	US-09-626-580C-64	Sequence 64, Appl
14	486.5	40.1	1085	3	US-09-277-716-15	Sequence 15, Appl
15	485.5	40.0	696	3	US-09-459-956-4	Sequence 4, Appli
16	485.5	40.0	720	3	US-09-839-650-1	Sequence 1, Appli
17	485.5	40.0	720	3	US-10-652-703A-1	Sequence 1, Appli
18	485.5	40.0	720	3	US-10-021-818A-3	Sequence 3, Appli
19	477.5	39.3	696	3	US-09-459-956-3	Sequence 3, Appli
20	474.5	39.1	1021	3	US-09-839-650-2	Sequence 2, Appli
21	474.5	39.1	1021	3	US-10-652-703A-2	Sequence 2, Appli
22	474.5	39.1	1104	3	US-09-277-716-30	Sequence 30, Appl
23	474.5	39.1	1104	3	US-09-609-161B-30	Sequence 30, Appl
24	474.5	39.1	1279	3	US-09-277-716-31	Sequence 31, Appl
25	474.5	39.1	1279	3	US-09-609-161B-31	Sequence 31, Appl
26	468.5	38.6	1482	3	US-09-977-897-1	Sequence 1, Appli
27	464.5	38.3	720	3	US-10-021-818A-1	Sequence 1, Appli
28	218.5	18.0	1559	3	US-09-049-475-6	Sequence 6, Appli
29	215.5	17.8	717	3	US-09-023-946B-3	Sequence 3, Appli
30	215.5	17.8	4196	3	US-09-453-313-1	Sequence 1, Appli
31	215.5	17.8	4199	3	US-09-504-117B-1	Sequence 1, Appli
32	215.5	17.8	6232	3	US-09-796-575-1	Sequence 1, Appli
33	215.5	17.8	7455	3	US-09-220-557-19	Sequence 19, Appl
34	215.5	17.8	7455	3	US-10-219-227-19	Sequence 19, Appl
35	215.5	17.8	7686	3	US-09-502-710-23	Sequence 23, Appl
36	215.5	17.8	7686	3	US-09-502-710-26	Sequence 26, Appl
37	215.5	17.8	7686	3	US-09-502-711-23	Sequence 23, Appl
38	215.5	17.8	7686	3	US-09-502-711-26	Sequence 26, Appl
39	215.5	17.8	7686	3	US-09-565-616A-2	Sequence 2, Appli
40	215.5	17.8	7687	3	US-09-502-710-24	Sequence 24, Appl
41	215.5	17.8	7687	3	US-08-818-604-31	Sequence 31, Appl
42	212.5	17.5	717	2	US-09-502-711-24	Sequence 24, Appl
43	212.5	17.5	717	3	US-09-346-946-31	Sequence 31, Appl
44	212.5	17.5	717	3	US-09-023-946B-21	Sequence 21, Appl
45	212.5	17.5	764	2	US-08-818-604-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1  
US-09-459-956-6  
; Sequence 6, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; TITLE OF INVENTION: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma BP  
US-09-459-956-6

Alignment Scores:  
Pred. No.: 3.24e-160  
Score: 1214.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 3  
Length: 678  
Matches: 225  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0



US-10-006-922A-12 (1-225) x US-09-459-956-6 (1-678)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
DB	1	ATGAGGCTCTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACGGTCAATGGGCACGAGTTTGAATGAAAGCGAAGAGAGAGGGGAGCCATACGAAGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACAATACCGTAAGCTTAAGGTAACCAAGGGGGGACCTTGGCATTGGCTTGGAATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
DB	181	TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
DB	241	GACTATATAAAGCTGTCTATTCTCGAAGGATTTAAATGGGAAGGGTCAAGAACTTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	301	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
DB	361	AAGGTCAAGTTCATTGGCGTGAACCTTTCCTCCGATGACCTGTATGCAAAAGAGACA	420
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
DB	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCCTCGTATGGCGGTGGAAGAGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
DB	481	ATTCAATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	541	TACATGGCAAGAAGCCTGTGACGCTAACGAGGTACTACTATGTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGACAGATATGAAGAACCAGAGGAGCG	660
QY	221	HisHisLeuPheLeu 225	
DB	661	CACCATCTGTCTCTT 675	

RESULT 2  
 US-09-866-538-11  
 ; Sequence 11, Application US/09866538  
 ; Patent No. 6852849  
 ; GENERAL INFORMATION:  
 ; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 ; APPLICANT: TSIEH, Roger  
 ; APPLICANT: Campbell, Robert  
 ; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
 ; FILE REFERENCE: REGEN1530-2  
 ; CURRENT APPLICATION NUMBER: US/09/866,538  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 11  
 ; LENGTH: 859  
 ; TYPE: DNA  
 ; ORGANISM: *Discozoma* sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (54) .. (731)  
 ; US-09-866-538-11

Alignment Scores:			
Pred. No.:	4.72e-160	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922A-12 (1-225) x US-09-866-538-11 (1-859)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
DB	54	ATGAGGCTCTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGGTTTCGCATGGAAGCA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	114	ACGGTCAATGGGCACGAGTTTGAATAAGGCGAAGAGAGAGGGGAGCCATACGAAGGC	173
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAATACCGTAAGCTTTAAGGTAACCAAGGGGGGACCTTTCATTGCTTGGAATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
DB	234	TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGCAACACCCTGCCACATACCA	293
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
DB	294	GACTATATAAAGCTGTCTATTCTCGAAGGATTTAAATGGGAAGGGTCATGAACCTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	354	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACAGATGGCTGTTTCATCTAC	413
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
DB	414	AAGGTCAAGTTCATTGGCGTGAACCTTTCCTCCGATGACCTGTATGCAAAAGAGACA	473
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
DB	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCCTCGTATGGCGGTGGAAGAGAG	533
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
DB	534	ATTCAATAAGGCTCTGAAGCTGAAGACGGTGGTCATTACCTAGTTGAATTCAAAAGTATT	593
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	594	TACATGGCAAGAAGCCTGTGACGCTAACGAGGTACTACTATGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGACAGATATGAAGAACCAGAGGAGCG	713
QY	221	HisHisLeuPheLeu 225	
DB	714	CACCATCTGTCTCTT 728	

RESULT 3  
 US-09-865-291-11  
 ; Sequence 11, Application US/09865291  
 ; Patent No. 6900304  
 ; GENERAL INFORMATION:  
 ; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 ; APPLICANT: TSIEH, Roger  
 ; APPLICANT: TING, Alice  
 ; APPLICANT: ZHANG, Jin  
 ; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION  
 ; FILE REFERENCE: REGEN1550  
 ; CURRENT APPLICATION NUMBER: US/09/865,291  
 ; CURRENT FILING DATE: 2001-05-24  
 ; NUMBER OF SEQ ID NOS: 42



```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
;
US-09-865-291-11

```

Alignment Scores:	
Pred. No.:	4.72e-160
Score:	1214.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	3
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
	859
	225
	0
	0
	0
	0

US-10-006-922A-12 (1-225) x US-09-865-291-11 (1-859)

QY	1	MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly	20
Db	54	ATGAGGCTTCCAAAGATGTTATCAAGGAGTTCAAGAGTTTAAAGTTCCGATGGAAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyIleArgProTyrGluGly	40
Db	114	ACGGTCATGGGCACGAGTTTGAATAGAAAGCGGAAGAGAGGAGGCCATTACGAAGGC	173
QY	41	HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	174	CACATAACCGTAAGCTTAAGGTAAACCAAGGGGACCCTTGCCATTGTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	234	TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA	293
QY	81	AspTyrLeuIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAGCTGTCAATTCCTGAAGATTAAATGGAAAGGTCATGAACCTTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACCGTGGCGTCGTTACTGTAAACCAAGATTCAGTTTGCAAGATGCTGTTCACTAC	413
QY	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr	140
Db	414	AAGGTCAAGTTCAATTGGCGTGAACTTTCCTCCGATGGAACTGTTATGCAAAAGAAGACA	473
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCACGACCTAGCGTTGTATCTCGTGATGGCGTGTGAAGGAGAG	533
QY	161	IleHisIysAlaLeuIysLeuIysAspGlyGlyHisIstYrLeuValGluPheIysSerIle	180
Db	534	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGTCACTTAACCTAGTTGAATTCAAAGTATT	593
QY	181	TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp	200
Db	594	TACATGGCAAGAACCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCCAACGAAGACTATACAATCGTTGAGCAGATATGAAGAACCAGAGGAGCG	713
QY	221	HisHisLeuPheLeu	225
Db	714	CACCATCTGTCTCTT	728

RESULT 4  
US-10-152-296-1  
; Sequence 1, Application US/10152296  
; Patent No. 6723537  
; GENERAL INFORMATION:

```

; APPLICANT: peelle, Beau
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells
; FILE REFERENCE: 021044-000110US
; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(723)
; OTHER INFORMATION: DsRED
;
US-10-152-296-1

```

Alignment Scores:	
Pred. No.:	1.31e-159
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	3
Length:	723
Matches:	224
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-10-152-296-1 (1-723)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	4	GTGCGCTCCTCCAGAAGCTCATCAAGAGATTATGCGCTTCAAGGTGCGCATGAGGGC	63
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGlyArgProTyrGluGly	40
Db	64	ACCTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC	123
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	124	CACAACACCGTGAAAGTGAAGGTGACCAAGGGCGGCCCCCTGCTTCGGCATCC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	184	CTGTCCCCCAGTTCAGTAGCGGCTCCAAGGTGATCGTGAAGCACCCCGCATCCCC	243
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	244	GACTACAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGAGCGCGTGATGAATTGAG	303
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	304	GACGGCGCGTGTGACCGTGAACCCAGSACTCTCCCTGCAGGACGGGTCTCATCTAC	363
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	364	AAGGTGAAGTTATCGCGCTGAACCTTCCCTCCGACGGCCCCGTAAATGCAGAGAAGACC	423
QY	141	MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	424	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCCGACGCGCGTGTGAAGGCGAG	483
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	484	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTGTGAGTTCAAGATATC	543
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	544	TACATGGCCAAGAAGCCCGTGCAGCTGCCCGCTACTACTACGTGACTCCAAGCTGGAC	603



```

QY      201  ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg  220
Db      604  ATCACTCCACACGAGACTACACCATGTGTGAGAGTAAGAGCGACCGAGGCGC      663

QY      221  HisHisLeuPheLeu  225
Db      664  CACCACCTGTTCTG  678

RESULT 5
US-10-001-189-45
; Sequence 45, Application US/10001189
; Patent No. 6962810
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 6984
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
US-10-001-189-45

Alignment Scores:
Pred. No.:      4.87e-158      Length:      6984
Score:          1210.00      Matches:      224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches:  0
Query Match:     99.67%      Indels:      0
DB:              3      Gaps:      0

US-10-006-922A-12 (1-225) x US-10-001-189-45 (1-6984)

QY      1  MetArgSerSerLyAsnValIleLySgluPheMetArgPheLySValArgMetGluGly  20
Db      2974  GTGCGCTCCTCCAAGACGTGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC  30333

QY      21  ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly  40
Db      3034  ACCGTGAACGGCCACGAGTTCCAGATCGAGGCGGAGGGCGGCCGCCCTACGAGGC  30933

QY      41  HisAsnThrValLySLeuLySValThrLySglYgLyProLeuProPheAlaTrpAspIle  60
Db      3094  CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCGCTTGAGACATC  31533

QY      61  LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro  80
Db      3154  CTGTCCCCCAGTTCAGTACGAGCTCCAAAGGTGTACGTGAAGCACCCCGCATCCCC  32133

QY      81  AspTyrLySLeuSerPheProGluGlyPheLySTrGluArgValMetAsnPheGlu  100
Db      3214  GACTACAAGAAGCTGTCTTCCCCAGGGCTTCAAGTGGAGCGGTGATGACTTCAG  32733

QY      101  AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr  120
Db      3274  GACGGCGGCGGTGTGACCGGTGACCCAGGACTCTCCCTGACAGACGGCTGTCTCATCTAC  333333

QY      121  LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr  140

```

```

Db      3334  AAGGTGAAGTTCATCGGCGGAACTTCCCTCCGACGGCCCCGGTATATGCAAGAAGACC 3339
               |||
QY      141  MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
               |||
Db      3394  ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGGTGCTGAAGCGCGAG 345
               |||
QY      161  IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
               |||
Db      3454  ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCACCTACTGCTGAGTTCAAGTCCATC 351
               |||
QY      181  TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
               |||
Db      3574  ATCACTCCACACAAGAGACTACACCATCGTGAGACGAGTACGAGCGCACCGAGCGCGC 363
               |||
QY      221  HisHisLeuPheLeu 225
               |||
Db      3634  CACCACCTGTTCTCTG 3648
               |||

RESULT 6
US-09-459-956-5
; Sequence 5, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; TITLE OF INVENTION: OPTICAL METHODS
; FILE REFERENCE: REGEM1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; CURRENT FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Discosoma striata
US-09-459-956-5

Alignment Scores:
Pred. No.:      2.87e-92      Length:      699
Score:          729.50      Matches:      133
Percent Similarity: 76.00%      Conservative: 38
Best Local Similarity: 59.11%      Mismatches:  53
Query Match:    60.09%      Indels:      1
DB:             3          Gaps:      1

US-10-006-922A-12 (1-225) x US-09-459-956-5 (1-699)

QY      1  MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
               |||
Db      1  ATGAGTGTGTTCCAAAGAGTGATCAAGCAAGAAATGTGATCGATCTTCATCTGGAAGCA 60
               |||
QY      21  ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
               |||
Db      61  ACGTTCAATGGGCACCTACTTGAATAATAAAGGCAAAAGGAAAGGACAGCCTAATGAAGGC 120
               |||
QY      41  HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
               |||
Db      121  ACCAATACCGTCACGCTCGAGGTTAACCAAGGGTGACCTCTGCCATTGTGTCATATT 180
               |||
QY      61  LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
               |||

```



Db 181 TTGTGCCCAATTTAGATGGAACAAGCATTGTCCACCACTTGACAAACATACAT 240  
 Qy 81 AspTyrLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPhgIu 100  
 Db 241 GATTATCTAAAGCTGTCAATTCGAGGAGATATACATGGACGGTCCATGCACCTTTGAA 300  
 Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
 Db 301 GACGGTGGCTGTGTGTATACCAATGATATCAGTTTGACAGGCACTGTTCTACTAC 360  
 Qy 121 LysValLysPheIleGlyValAsnPhgProSerAspGlyProValMetGlnLysLysThr 140  
 Db 361 GACATCAAGTTCACTGGCTTGAACTTCTCCAAATGACCCGTTGTGCAGAAAGACACA 420  
 Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
 Db 421 ACTGGCTGGAAACCGACACTGAGCGTTGTATCTCGTATGGTGTGTATAGAGAC 480  
 Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
 Db 481 ATCCATCATGTCTGACAGTTGAAGAGGTGTCATTACGCATGTGACATTAAACTGTT 540  
 Qy 181 TyrMetAlaLysLys---ProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu 199  
 Db 541 TACAGGGCCCAAGAGCGCGCTTGAAAGATGCCAGGTATCACTATGTGACACCAACTG 600  
 Qy 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 219  
 Db 601 GTTATATGGAACAACGACAAGAAATTCAATAAGTTGAGAGCATGAATACTGCCGTTGCA 660  
 Qy 220 ArgHisHisLeuPhe 224  
 Db 661 CGCCACCATCCGTTCC 675

RESULT 7

US-10-244-779-1  
 ; Sequence 1, Application US/10244779  
 ; Patent No. 6933375  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falkowski, Paul  
 ; APPLICANT: Sun, Yi  
 ; APPLICANT: Gorbunov, Maxim  
 ; APPLICANT: Wyman, Kevin  
 ; APPLICANT: Chen, Yi-Bu  
 ; TITLE OF INVENTION: mCFP Encoding Nucleic Acids,  
 ; TITLE OF INVENTION: Polypeptides, Antibodies and Methods of Use Thereof  
 ; FILE REFERENCE: Rut 00-0023US  
 ; CURRENT APPLICATION NUMBER: US/10/244,779  
 ; PRIOR FILING DATE: 2002-12-19  
 ; PRIOR APPLICATION NUMBER: 60/322,189  
 ; PRIOR FILING DATE: 2001-09-14  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 860  
 ; TYPE: DNA  
 ; ORGANISM: Montastrea cavernosa  
 ; US-10-244-779-1

Alignment Scores:

Pred. No.: 2.52e-85 Length: 860  
 Score: 681.00 Matches: 123  
 Percent Similarity: 73.95% Conservative: 36  
 Best Local Similarity: 57.21% Mismatches: 56  
 Query Match: 56.10% Indels: 0  
 DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-244-779-1 (1-860)

Qy 6 AsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHis 25  
 Db 138 AGTGTGATTAAATCAGTCATGAATAATCAAGCTCGTATGACGGCATTTGTAAACGGGAC 197

Qy 26 GluPheGluIleGluGlyGluGluGlyLysArgProTyrGluGlyHisAsnThrValLys 45  
 Db 198 AAGTTCATGATTACAGGAGAGAGGTGAAGGCAAGCCCTTTCAGGGGAACACACACTAATA 257  
 Qy 46 LeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGlnPhe 65  
 Db 258 CTTAAGTCAAGAAGGCGGACTCTGCTTTCGCTTACGACATCTTGACACACAGCATTT 317  
 Qy 66 GlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeu 85  
 Db 318 CAGTACGGCAACAGGATATTCACCAATAATCCAAAAGACATACCACTATTCAAGCAG 377  
 Qy 86 SerPheProGluGlyPheLysTyrGluArgValMetAsnPhgIuAspGlyValVal 105  
 Db 378 TCGTTCTGAGGGGTATTCCTGGGAAAGAACATGACTTTCGAAGACAGGCGTTGC 437  
 Qy 106 ThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIle 125  
 Db 438 ACCGTCAACAGCGACATTAAGTTGGAAGCGGACTGTTTTCACGAATTCGATTTTAT 497  
 Qy 126 GlyValAsnPhgProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGluAla 145  
 Db 498 GGTGTGAACCTTCCCTCCAGTGTTCAGTTATGACAGAGAGACGCTGAATGGAGCCA 557  
 Qy 146 SerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeu 165  
 Db 558 TCCACTGGAATATGTACGTGCGTGAATGAGTGTACTGGGGGATGTTAGCAGACGCTG 617  
 Qy 166 LysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLys 185  
 Db 618 TTGCTGAAGGGGATAAATCATCACCGATGTAACTTCAGAACTTACGGGGCAAGAAG 677  
 Qy 186 ProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsn 205  
 Db 678 GGTGTGCTGTTGCCAGATATCACTTGTGACCAACGAATGAATCTGAGCCATGAC 737  
 Qy 206 GluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
 Db 738 AAAGATTACAACACCGTTGAGGTGTATGAGAATGCCGTTGCTCCG 782

RESULT 8

US-09-459-956-7  
 ; Sequence 7, Application US/09459956  
 ; Patent No. 6342379  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tsien, Roger Y.  
 ; APPLICANT: Gonzalez, Ili, Jesus E.  
 ; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
 ; TITLE OF INVENTION: OPTICAL METHODS  
 ; FILE REFERENCE: REGEN1290-4  
 ; CURRENT APPLICATION NUMBER: US/09/459,956  
 ; PRIOR FILING DATE: 1999-12-13  
 ; PRIOR APPLICATION NUMBER: 08/765,860  
 ; PRIOR FILING DATE: 1999-05-08  
 ; PRIOR APPLICATION NUMBER: 08/481,977  
 ; PRIOR FILING DATE: 1995-06-07  
 ; PRIOR APPLICATION NUMBER: PCT/US96/09652  
 ; PRIOR FILING DATE: 1996-06-06  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 801  
 ; TYPE: DNA  
 ; ORGANISM: Clavularia sp  
 ; US-09-459-956-7

Alignment Scores:

Pred. No.: 2.06e-83 Length: 801  
 Score: 667.00 Matches: 121  
 Percent Similarity: 73.76% Conservative: 42  
 Best Local Similarity: 54.75% Mismatches: 58  
 Query Match: 54.94% Indels: 0  
 DB: 3 Gaps: 0



US-10-006-922A-12 (1-225) x US-09-459-956-7 (1-801)

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22  
Db 121 ACCACCATTGGGTGATTAAACCAGACATGAAGATTAAAGCTGAAGATGGAAGAAATGTA 180  
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsn 42  
Db 181 AACGGGCATGCTTTGTGATCGAAGGAGAGAGAAAGAAAGCCTTACGATGGGACACAC 240  
QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62  
Db 241 ACTTTAACTCGAAGTGAAGGAGTGGCCTCTGCCTTTTCTTACGATATCTGTCA 300  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
Db 301 AACGCGTCCAGTACGGAACAGACATGACAAATACCCAGACGATATAGCAGACTAT 360  
QY 83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102  
Db 361 TTCAAGCAGTCGTTTCCCGAGGGATATTCCTGGGAAAGAACCATGACTTTGAAAGACAA 420  
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
Db 421 GGCATTGTCAAAAGTGAAGAAAGTACATAGCATGAGAGAAAGACTCTTTATCTATGAAT 480  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
Db 481 CGTTTGATGGAGTAACTTTCCTCCCAATGCTCCGTTATGCAAAAAAACTTTGAAG 540  
QY 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162  
Db 541 TGGGAACCATCCACTGAGATTATGTACGTGCGTGATGAGTGTGTCGGAATATTAGC 600  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
Db 601 CATCTCTCTGCTGAGAGGAGGTGGCCATTACCGATGTGACTTCAAAAGTATTACAA 660  
QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
Db 661 GCAAAAAAGTTGTCAAAATTGCCAGACTATCACTTGTGACCATTCGCAATTGAGATCTTG 720  
QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222  
Db 721 AACCATGCAAGGATTACAAACAAGTAACGCTGTATGAAATGCAAGTGTCTGCTATTCT 780  
QY 223 Leu 223  
Db 781 TTG 783  
RESULT 9  
US-09-459-956-2  
; Sequence 2, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FaabSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 690

; TYPE: DNA  
; ORGANISM: Anemonia majano  
US-09-459-956-2  
Alignment Scores:  
Pred. No.: 2,11e-59  
Score: 495.00  
Percent Similarity: 64.68%  
Best Local Similarity: 45.77%  
Query Match: 40.77%  
DB: 3  
Matches: 690  
Conservative: 92  
Mismatches: 38  
Indels: 2  
Gaps: 1

US-10-006-922A-12 (1-225) x US-09-459-956-2 (1-690)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGCTCTTCAACAAGTTTATCGAGATGACATGAATAAGCCTTACCATATGATGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 TGTGTCAATGGGCATTACTTTAACCGTCAAAGGTGAAGCAACGGGAAGCCATAGAGGG 120  
QY 41 HisAsnThrValLysLeuLysValThr-----LysGlyGlyProLeuProPheAlaTrp 58  
Db 121 ACCGAGACTTCGACTTTTAAAGTCACCATGGCCAAACGGTGGGCCCTTGCAATCTCCTTT 180  
QY 59 AspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAsp 78  
Db 181 GACATACTATCTACAGTGTCAAAATATGGAATGCAATGCTTACTCGGTATCTTACCAGT 240  
QY 79 IleProAspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsn 98  
Db 241 ATGCCGCACTATTTCAAACAAGCATTTCTGACCGAATGTCTATATGAAGACTTTTACC 300  
QY 99 PheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe 118  
Db 301 TATGAAGATGAGAGTGTCTACAGCCAGTTGGGAATGAAGCCTTAAAGGCAACTGCTTT 360  
QY 119 IleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLys 138  
Db 361 GAGCACAATCCACGTTTCATGAGTGAACCTTCTGCTGATGAGCCTGTGATGGCAAG 420  
QY 139 LysThrMetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLys 158  
Db 421 AAGACAACGTGTTGGACCACATCTTTGAGAAATGACTGTCTGCGATGGAATATTGAAG 480  
QY 159 GlyGluIleHisLysAlaLeuLysLysAspGlyGlyHisTyrLeuValGluPheLys 178  
Db 481 GGTGATGTCAACCGCTTCTCATGCTGCAAGGAGGTGGCAATTACAGATGCCAATTCAC 540  
QY 179 SerIleTyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLys 198  
Db 541 ACTTCTTACAGACAAAAAACCGGTGACGATGCCAACCAACCATGTGTGGAACATCGC 600  
QY 199 Leu 199  
Db 601 ATT 603  
RESULT 10  
US-09-609-161B-15  
; Sequence 15, Application US/09609161B  
; Patent No. 6436682  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUCI  
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH  
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS  
; FILE REFERENCE: 24729-121B  
; CURRENT APPLICATION NUMBER: US/09/609,161B  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/277,716



/ PRIOR FILING DATE: 1999-03-26  
/ PRIOR APPLICATION NUMBER: 60/102,939  
/ PRIOR FILING DATE: 1998-10-01  
/ PRIOR APPLICATION NUMBER: 60/089,367  
/ PRIOR FILING DATE: 1998-06-15  
/ PRIOR APPLICATION NUMBER: 60/079,624  
/ PRIOR FILING DATE: 1998-03-27  
/ NUMBER OF SEQ ID NOS: 32  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 15  
/ LENGTH: 1079  
/ TYPE: DNA  
/ ORGANISM: Renilla mulleri  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (259)..(975)  
/ OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)  
US-09-609-161B-15

Alignment Scores:  
Pred. No.: 6.71e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-609-161B-15 (1-1079)

QY 2 ArgSerSerLySAAnValIleLyS-----GluPheMetArgPheLySVal 16  
DB 256 AAGATGAGTAAACAATATGTAAGAACAACACTTGTTTACAAGAGTAATGCGTATAAAGTA 315  
QY 17 ArgMetGluGlyThrValAaNGlyHisGluPheGluIleGluGlyGluGlyArg 36  
DB 316 AATCTGGAAGGAAATGTAAACAACCATGTTTACAATGGAAGGTTGCGGCAAGGGAAT 375  
QY 37 ProTyrgLugLyHisAAnthrValLySLeuLySValThrLySGLyGlyProLeuProPhe 56  
DB 376 ATTTTATTCGGCAATCAACTGCTTCAGATTCGTGTCAAGAAAGGGCCCCACTGCTTTT 435  
QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyrgLySerLySValTyrgValLyHisPro 76  
DB 436 GCATTTGATATTGTGTCAACCACTTTTCAATATGGCAACCGTACTTTCACGAATATCCG 495  
QY 77 AlaAspIleProAspTyrgLySLeuSerPheProGluGlyPheLySTrgLiuArgVal 96  
DB 496 AATGATATATCAGATTATTATTATACAATCATTTCCAGCAGGATTATGTATGAACGAACA 555  
QY 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGly 116  
DB 556 TTACGTTACGAAGATGGCGGACTTGTGAATTCGTTCAAGATATAAATTTAATAGAAGAC 615  
QY 117 CysPheIleTyrgLySValLySLeuSerPheIleGlyValAsnPheProSerAspGlyProValMet 136  
DB 616 AAGTTGCTTACAGAGTGGGAATACAAAGGTAGTAACTTCCAGATGATGTCGCCGTATG 675  
QY 137 GlnLySlySThrMetGlyTrgLiuAlaSerThrGluArgLeuTyrgProArgAspGlyVal 156  
DB 676 CAGAAAGACTATCTTAGAATAGAGCCCTTCATTGAAAGCCATGTAATGAATATGGCGTC 735  
QY 157 LeuLySGLyGluIleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrgLeuValGlu 176  
DB 736 TTGGTCGGCGAAGTAATCTTGTCTATAAACTAACTCTGGGAATATATTATTCATGTAC 795  
QY 177 PheLySerIleTyrgMetAlaLySLeuSProVal---GlnLeuProGlyTyrgTyrgVal 195  
DB 796 ATGAACAACATTAATGAAGTCGAAGGTGTAGTAAGAAGATTTCCTTCGTATCATTTTAT 855  
QY 196 AspSerLySLeuAspIleThrSerHisAaNGlyAspTyrgThrIleValGluGlnTyrgLiu 215  
DB 856 CAACATCGTTTGGAAAAGACT--TACGTAGAAGACGGGGGTTTCGTTGAACAGCATAGAG 912

RESULT 11  
US-09-626-581D-64  
/ Sequence 64, Application US/09626581D  
/ Patent No. 6548249  
/ GENERAL INFORMATION:  
/ APPLICANT: Anderson, David  
/ TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
/ TITLE OF INVENTION: Libraries  
/ FILE REFERENCE: A-66900-3/RMS  
/ CURRENT APPLICATION NUMBER: US/09/626,581D  
/ PRIOR FILING DATE: 2000-07-27  
/ PRIOR APPLICATION NUMBER: 09/169,015  
/ PRIOR FILING DATE: 1998-10-08  
/ PRIOR APPLICATION NUMBER: 09/415,765  
/ PRIOR FILING DATE: 1999-10-08  
/ NUMBER OF SEQ ID NOS: 65  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 64  
/ LENGTH: 1079  
/ TYPE: DNA  
/ ORGANISM: Renilla muelleri  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (259)..(975)  
/ OTHER INFORMATION:  
US-09-626-581D-64

Alignment Scores:  
Pred. No.: 6.71e-58 Length: 1079  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-626-581D-64 (1-1079)

QY 2 ArgSerSerLySAAnValIleLyS-----GluPheMetArgPheLySVal 16  
DB 256 AAGATGAGTAAACAATATGTAAGAACAACACTTGTTTACAAGAGTAATGCGTATAAAGTA 315  
QY 17 ArgMetGluGlyThrValAaNGlyHisGluPheGluIleGluGlyGluGlyArg 36  
DB 316 AATCTGGAAGGAAATGTAAACAACCATGTTTACAATGGAAGGTTGCGGCAAGGGAAT 375  
QY 37 ProTyrgLugLyHisAAnthrValLySLeuLySValThrLySGLyGlyProLeuProPhe 56  
DB 376 ATTTTATTCGGCAATCAACTGCTTCAGATTCGTGTCAAGAAAGGGCCCCACTGCTTTT 435  
QY 57 AlaTrpAspIleLeuSerProGlnPheGlnTyrgLySerLySValTyrgValLyHisPro 76  
DB 436 GCATTTGATATTGTGTCAACCACTTTTCAATATGGCAACCGTACTTTCACGAATATCCG 495  
QY 77 AlaAspIleProAspTyrgLySLeuSerPheProGluGlyPheLySTrgLiuArgVal 96  
DB 496 AATGATATATCAGATTATTATTATACAATCATTTCCAGCAGGATTATGTATGAACGAACA 555  
QY 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGly 116  
DB 556 TTACGTTACGAAGATGGCGGACTTGTGAATTCGTTCAAGATATAAATTTAATAGAAGAC 615  
QY 117 CysPheIleTyrgLySValLySLeuSerPheIleGlyValAsnPheProSerAspGlyProValMet 136  
DB 616 AAGTTGCTTACAGAGTGGGAATACAAAGGTAGTAACTTCCAGATGATGTCGCCGTATG 675  
QY 137 GlnLySlySThrMetGlyTrgLiuAlaSerThrGluArgLeuTyrgProArgAspGlyVal 156  
DB 676 CAGAAAGACTATCTTAGAATAGAGCCCTTCATTGAAAGCCATGTAATGAATATGGCGTC 735  
QY 157 LeuLySGLyGluIleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrgLeuValGlu 176  
DB 736 TTGGTCGGCGAAGTAATCTTGTCTATAAACTAACTCTGGGAATATATTATTCATGTAC 795







Db 556 TTACGTTACGAAGATGGCGGACTTGTGAATTCGTTCAAGATATATAATTATAGAGAC 615  
Qy 117 CysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMet 136  
Db 616 AAGTTCGTCTACAGAGTGAATACAAAGTAGTACTTCCAGATGATGGTCCCGTCATG 675  
Qy 137 GlnLysLysThrMetGlyTyrGluAlaSerThrGluArgLysTyrProArgAspGlyVal 156  
Db 676 CAGAGACTAATCTTAGAATAGACCTTCACTTGAAGCCATGATCATGAATAATGCGGTC 735  
Qy 157 LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlu 176  
Db 736 TTGTCGCGCAAGTAATTCTGTCTATAAATACTCTGGAAATATTAATTCATGTCAC 795  
Qy 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrTyrVal 195  
Db 796 ATGAAAACATTAAATGAAGTCGAAGGTGTAGTAAGAGTTCCCTCGATCATTTTATT 855  
Qy 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215  
Db 856 CAACATCGTTGGAAAAAGACT---TACGTAGAAGACGGGGGTTCGTTGAACACGATGAG 912

RESULT 14

US-09-277-716-15  
; Sequence 15, Application US/09277716A  
; Patent No. 6232107  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  
; CURRENT APPLICATION NUMBER: US/09/277,716A  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/102,939  
; EARLIER FILING DATE: 1998-10-01  
; EARLIER APPLICATION NUMBER: 60/089,367  
; EARLIER FILING DATE: 1998-06-15  
; EARLIER APPLICATION NUMBER: 60/079,624  
; EARLIER FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Renilla mulleri  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(975)  
; FEATURE:  
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)  
US-09-277-716-15

Alignment Scores:

Pred. No.: 6.77e-58 Length: 1085  
Score: 486.50 Matches: 97  
Percent Similarity: 64.09% Conservative: 44  
Best Local Similarity: 44.09% Mismatches: 72  
Query Match: 40.07% Indels: 7  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-277-716-15 (1-1085)

Qy 2 ArgSerSerLysAsnValIleLys-----GluPheMetArgPheLysVal 16  
Db 256 AAGATGAGTAAACAATATGGAAGAACACTTGTTTACAGAAAGTAATGTCGTATAAGTA 315  
Qy 17 ArgMetGluGlyThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArg 36  
Db 316 AATCTGAAGGAATTTGTAACACCATGTTTACAAATGAGAGGTTGCGCAAGGAAT 375  
Qy 37 ProTyrGluGlyHisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPhe 56

Db 376 ATTTATTCGGCAATCACTGTTCAAGATTCGTGTACGAAGGGCCCACTGCTTTT 435  
Qy 57 AlaTyrAspIleLeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPro 76  
Db 436 GCATTGTGATATTGTGTCAACCAGCTTTTCAATATGGCAACCGTACTTTCACGAATATCCG 495  
Qy 77 AlaAspIleProAspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgVal 96  
Db 496 AATGATATATCAGATTATTATTATACATCAATTCATTCCAGCAGAGATTATGTGAACGAACA 555  
Qy 97 MetAsnPheGluAspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGly 116  
Db 556 TTACGTTACGAAGATGCGGACTTGTGAATTCGTTCAAGATTAATTAATAGAGAC 615  
Qy 117 CysPheIleTyrLysValLysPheIleGlyValAsnPheProSerAspGlyProValMet 136  
Db 616 AAGTTCGTCTACAGAGTGAATACAAAGTAGTAAGTCCCAAGATGATGTCGCCGTCATG 675  
Qy 137 GlnLysLysThrMetGlyTyrGluAlaSerThrGluArgLysTyrProArgAspGlyVal 156  
Db 676 CAGAGACTAATCTTAGAATAGAGCCTTCATTTGAAGCCATGTACATGAATAATGCGGTC 735  
Qy 157 LeuLysGlyGluIleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlu 176  
Db 736 TTGTCGCGCAAGTAATTCTGTCTATAAATACTCTGGAAATATTAATTCATGTCAC 795  
Qy 177 PheLysSerIleTyrMetAlaLysLysProVal---GlnLeuProGlyTyrTyrTyrVal 195  
Db 796 ATGAAAACATTAAATGAAGTCGAAGGTGTAGTAAGAGTTCCCTCGATCATTTTATT 855  
Qy 196 AspSerLysLeuAspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlu 215  
Db 856 CAACATCGTTGGAAAAAGACT---TACGTAGAAGACGGGGGTTCGTTGAACACGATGAG 912

RESULT 15

US-09-459-956-4  
; Sequence 4, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; EARLIER FILING DATE: 1999-12-13  
; EARLIER APPLICATION NUMBER: 08/765,860  
; EARLIER FILING DATE: 1999-05-08  
; EARLIER APPLICATION NUMBER: 08/481,977  
; EARLIER FILING DATE: 1995-06-07  
; EARLIER APPLICATION NUMBER: PCT/US96/09652  
; EARLIER FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Zoanthus sp  
US-09-459-956-4

Alignment Scores:

Pred. No.: 4.6e-58 Length: 696  
Score: 485.50 Matches: 95  
Percent Similarity: 64.71% Conservative: 37  
Best Local Similarity: 46.57% Mismatches: 67  
Query Match: 39.99% Indels: 5  
DB: 3 Gaps: 3

US-10-006-922A-12 (1-225) x US-09-459-956-4 (1-696)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGCTCATTCAAAGCAGCGTCTAAAGAGAAATGACAATGAATAACACATGGAAGG 60



OY		21	ThrValAsnGlyHisGlutPheGluIleGlnGlyGlnGlyArgProTyrGlnGly	40
Dd		61	TGCCCAACGGAGCATAAATTGTGATCAGCGGCAGAAGCCATTGGATATCCGTTCAAAGGG	120
OY		41	HIsanThrValLysLeuLysValThrylsgLyglyProLeuProPhelaTrpaspile	60
Dd		121	AACAAGACTATACTGTGTGTGATCGAAGGGGACCATTGCCATTTCGAAGACATA	180
OY		61	LeuSerProglInphegIntyrglyserlysValTyrrvalylshisproalaaspilePro	80
Dd		181	TTGTCACTGGCTTTAAGTAGCACGAGACAGATTTCACCTGAATATCCTCAAGACATAGTA	240
OY		81	ASPtyrLysLysLeuSerpheProglInglyrhelystrpgluargvalmelasnphleglu	100
Dd		241	GACTATTCCAAGAACTCGTGCTCGTGATATACATGGGGCAGGCTTTCTTTGAG	300
OY		101	ASpGLgLYValValThrValThrGlnasp-----SerrseuGlnaspGlyCySpHe	118
Dd		301	GATGGAAGCAGTCTGCATATGCAATGTAGATATAACAGTGAGTGTCAAGAAAAC TGcAtt	360
OY		119	IletyrlsvallyspheilleglyValasnpheproserraspilyprovalmetGlnlys	138
Dd		361	TATCATTAAGACATATTTAATGAATGAATTTCTCTGCTGATGGA CCTIGATGA AAAAG	420
OY		139	LysThrmelGlyTrpGluAlaserthrGlnarginleu-----TyrProArGaspglyVal	156
Dd		421	ATGACAACTAAC TGGGAAGCAT C CTGGAG A AGAT CAT GCCAGT ACCT AA G CA GGSG ATA	480
OY		157	Leulysglygluilehislysalaleullysleullybaspglyghlytrylleuvaglu	176
Dd		481	CTGAAGGGGATGTC TC CAT GTAC CTC TTCT GAAG ATGTGTGG CGTT ACCG GT GCAG	540
OY		177	PheLysSerIleTyrMetAlalyslys---ProValGlnleuProGlyTyrtYrTYrrVal	195
Dd		541	TTCGACACAGTTTACAALAGCALAGCTGTGCCAAGTAAGATGCCGAGTGGCACTTCATC	600
OY		196	AspSerLysleu 199	
Dd		601	CAGCATTAAGCTC 612	

Search completed: January 12, 2006, 12:26:40  
Job time : 185 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 10:08:26 ; Search time 793 Seconds  
(without alignments)  
2346.293 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTYEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US10006922/runat\_10012006\_162409\_18977/app\_query.fasta\_1.391  
-DB=published\_Applications\_NA\_Main -QFMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922 @CGN\_1.1.1549 @runat\_10012006\_162409\_18977 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA\_Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	3	US-09-967-772-6
2	1214	100.0	678	5	US-10-006-922-11
3	1214	100.0	678	5	US-10-081-864-7
4	1214	100.0	678	5	US-10-121-258-2
5	1214	100.0	678	6	US-10-315-920-1
6	1214	100.0	678	6	US-10-132-067-3
7	1214	100.0	678	6	US-10-335-517-6

8	1214	100.0	678	6	US-10-334-288-6	Sequence 6, Appli
9	1214	100.0	678	7	US-10-311-030-5	Sequence 5, Appli
10	1214	100.0	678	8	US-10-656-029-21	Sequence 21, Appl
11	1214	100.0	678	9	US-10-505-486-27	Sequence 27, Appl
12	1214	100.0	678	9	US-10-844-064A-1	Sequence 1, Appli
13	1214	100.0	678	9	US-10-931-304-2	Sequence 2, Appli
14	1214	100.0	859	3	US-09-999-745-66	Sequence 66, Appl
15	1214	100.0	859	3	US-09-866-538-11	Sequence 11, Appl
16	1214	100.0	859	3	US-09-794-308-11	Sequence 11, Appl
17	1214	100.0	859	3	US-09-865-291-11	Sequence 11, Appl
18	1214	100.0	859	7	US-10-433-640-12	Sequence 12, Appl
19	1214	100.0	859	8	US-10-885-988-11	Sequence 11, Appl
20	1214	100.0	859	8	US-10-857-622-11	Sequence 11, Appl
21	1214	100.0	3311	3	US-09-797-496B-3	Sequence 3, Appli
22	1210	99.7	681	5	US-10-006-922-35	Sequence 35, Appl
23	1210	99.7	681	5	US-10-121-258-3	Sequence 3, Appli
24	1210	99.7	681	5	US-10-121-258-23	Sequence 23, Appl
25	1210	99.7	681	7	US-10-311-030-8	Sequence 8, Appli
26	1210	99.7	681	9	US-10-931-304-3	Sequence 9, Appli
27	1210	99.7	681	9	US-10-931-304-23	Sequence 23, Appl
28	1210	99.7	713	7	US-10-311-030-11	Sequence 11, Appl
29	1210	99.7	713	7	US-10-311-030-12	Sequence 12, Appl
30	1210	99.7	723	5	US-10-152-296-1	Sequence 1, Appli
31	1210	99.7	723	7	US-10-739-656-1	Sequence 1, Appli
32	1210	99.7	1638	5	US-10-214-932-51	Sequence 51, Appl
33	1210	99.7	1647	5	US-10-214-932-75	Sequence 75, Appl
34	1210	99.7	4692	6	US-10-161-403-29	Sequence 29, Appl
35	1210	99.7	4692	7	US-10-433-640-16	Sequence 16, Appl
36	1210	99.7	4692	10	US-11-006-076-29	Sequence 29, Appl
37	1210	99.7	5436	9	US-10-169-050-46	Sequence 46, Appl
38	1210	99.7	6984	5	US-10-001-189-45	Sequence 45, Appl
39	1210	99.7	7910	9	US-10-169-050-20	Sequence 20, Appl
40	1210	99.7	9320	8	US-10-609-019-4	Sequence 4, Appli
41	1210	99.7	9658	8	US-10-609-019-4	Sequence 4, Appli
42	1210	99.7	9678	8	US-10-609-019-3	Sequence 3, Appli
43	1210	99.7	10263	8	US-10-609-019-2	Sequence 2, Appli
44	1207	99.4	678	6	US-10-315-920-3	Sequence 3, Appli
45	1203	99.1	7508	7	US-10-742-828-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-967-772-6  
; Sequence 6, Application US/09967772  
; Patent No. US20020164577A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSIEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/09/967,772  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-09-967-772-6  
Alignment Scores: 6.83e-143 Length: 678  
Pred. No.: 1214.00 Matches: 225  
Score:



Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-967-772-6 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCCTCCAGAATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGCAGCAGTTGAAATAGAAAGCGAAGGAGAGGGGCAATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACATATACCGTAAGCTTAAGTAACCAAGGGGGACCTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCACAAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAAGCTGTCAATTCTCGAAGGATTAAATGGGAAGGGTCATGAACTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTGACTGTAAACCCAGGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTCATGGCGTGAACCTTCCGATGGACCTGTATGCAAAAGAAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCCTCGATGGCGTGTGAAGAAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCAAAAGAGCCTGTGCAGCTACAGAGGTACTATATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAAGCAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCCTT 675

RESULT 2  
US-10-006-922-11  
; Sequence 11, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898

; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/457,556  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/444,338  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma species  
US-10-006-922-11  
  
Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-006-922-11 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCCTCCAGAATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGCAGCAGTTGAAATAGAAAGCGAAGGAGAGGGGCAATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACATATACCGTAAGCTTAAGTAACCAAGGGGGACCTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCACAAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAAGCTGTCAATTCTCGAAGGATTAAATGGGAAGGGTCATGAACTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTGTACTGTAAACCCAGGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTCATGGCGTGAACCTTCCGATGGACCTGTATGCAAAAGAAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCCTCGTATGGCGGTGGAAGAAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCAAAAGAGCCTGTGCAGCTACAGGGTACTACTATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAAGCAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225



Db 661 CACCATCTGTTCTT 675

RESULT 3  
US-10-081-864-7  
; Sequence 7, Application US/10081864  
; Publication No. US20030022287A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Savitsky, Alexandr  
; APPLICANT: Fradkov, Arcady  
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-067  
; CURRENT APPLICATION NUMBER: US/10/081,864  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/270,983  
; PRIOR FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp  
; US-10-081-864-7

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-081-864-7 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGCGCTCTCCAGAAAGCTGAAGGTGACCAAGGCGGCCCTTCGCTTCAAGGTGCGCATGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCTTCGCTTCAAGGTGCGCATGAGGGC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTATGTAAGCAACCCCGCGCATATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACAAAGACTGTCTTCTCCCGAGGGCTTCAAGTGGGAGCGCGTGAATGAATTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGCGGTGTGACCGTGAACCCAGACTCTCTCTGACAGACGGCTGCTTCAATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCGGTGATGACAGAAAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGGAGCGCGGTGTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAGAAGACCCGTGACAGCTGCCCGGCTACTACTAGTGAAGTCCAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCACAAAGAGACTACACCATCTGTGAGAGACTACGAGCGCACCGAGGGCCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACTGTTCTTG 675

RESULT 4  
US-10-121-258-2  
; Sequence 2, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CPI  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(678)  
; OTHER INFORMATION: wild-type DsRed  
; US-10-121-258-2

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-2 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAGATGTATCAAGAGTTCATGAGGTTTAAGTTTGAAGTTCGATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCAAGATTGAATATGAAGCGCAAGAGAGAGGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAAGCTTAAGTAAACCAAGGGGGACCTTGGCCATTGCTTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTACACCACAATTCAGTATGAAGCAAGTATATGTCAAACACCTTGCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATATAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120



Db 301 GACGGTGGCGTCTGTTACTGTGTAACCCAGGATTCAGTTTGACGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGAAGCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGATGAGCGGTGTTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTGTGAAGCTGAAGAAGCGGTGTCATTACTAGTTGAATTCAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCGTGTGACCTACAGGGTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATACCAAGCCACAACGAGACTATACAATCGTTGAGCAGATGTAAAGAACCGAGGACGCC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTT 675

RESULT 5  
US-10-315-920-1  
; Sequence 1, Application US/10315920  
; Publication No. US20030175809A1  
; GENERAL INFORMATION:  
; APPLICANT: Fradkov, Arcady Fedorovich  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
; FILE REFERENCE: CLON-077CIP  
; CURRENT APPLICATION NUMBER: US/10/315, 920  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/211, 607  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: PCT/US01/19097  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant of sequence from Discosoma sp.  
US-10-315-920-1

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-315-920-1 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGCGCTCCTCCAAGAACGTTCATCAAGAGATTGCGCTTCAAGGTGCGCATGGAAGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGCGCACGAGTTGAGATCGAGGCGAGGGCGAGGGCCGCCCTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
4

Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCCCCCTTGCCCTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCGCATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300  
QY 301 GACGGCGCGGTGTAACCGTGACCCAGAGCTTCTCCCTGCAAGAACGGCTGCTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCGACGCGCCCGTGATGCAAGAAAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCAGCGCGGTGTAAGGGCAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGCGCCGTGAAGCTGAAGAGCGGCGCACTACCTGTTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCCAAGAACCCGTGACGCTGCCGCTACTACTAGTGAGTCAAGTCCAGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACAACGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTG 675

RESULT 6  
US-10-132-067-3  
; Sequence 3, Application US/10132067  
; Publication No. US20030203355A1  
; GENERAL INFORMATION:  
; APPLICANT: Bradbury, Andrew  
; APPLICANT: Zeytun, Ahmet  
; APPLICANT: Waldo, Geoffrey  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: Fluorobodies: Binding ligands with intrinsic  
; FILE REFERENCE: 021362-00600US  
; CURRENT APPLICATION NUMBER: US/10/132, 067  
; CURRENT FILING DATE: 2002-04-24  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; OTHER INFORMATION: red fluorescent protein (dsRED)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(678)  
; OTHER INFORMATION: dsRED  
US-10-132-067-3

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0



US-10-006-922A-12 (1-225) x US-10-132-067-3 (1-678)

QY 1 MetArgSerSerLySAsnValIleLysGluPheMetArgPheLySValArgMetGluGly 20  
1 ATGAGGCTCTTCCAAGAAATGTTATCAAGAGATTCAAGAGTTTAAGTTTAAGTTTCCGATGGAAAGCA 60  
DB 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
61 ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGAGAGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGLYGLYProLeuProPheAlaTrpAspIle 60  
121 CACAATAACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTGGCATTTGCTTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
181 TTGTCAACCACAATTTCAAGTATGGAAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
241 GACTATATAAAAGCTGTATTCCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
301 GACGGTGGCGTCTTACTGTAAACCCAGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360  
DB 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
361 AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGAGACCTGTTATGCAAAAGAAAGACA 420  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLYGLU 160  
421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAGAGAGAG 480  
QY 161 IleHisValAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuThr 180  
481 ATTCATTAAGGCTCTGAAGCTGAAAGACGGTGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
DB 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
541 TACATGGCAAGAACCTGTGACAGTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
601 ATAACAAGCCACAAGAACTATACAAATCGTTGAGACAGTATGAAGAACCAGAGGACGC 660  
DB 221 HisHisLeuPheLeu 225  
661 CACCATCTGTTCTT 675

RESULT 7  
US-10-335-517-6  
Sequence 6, Application US/10335517  
Publication No. US20030207248A1

GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

APPLICANT: TSJEN, Roger

APPLICANT: GONZALEZ, Jesus

TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS

FILE REFERENCE: REGEN1290-5

CURRENT APPLICATION NUMBER: US/10/335,517

PRIOR FILING DATE: 2002-12-31

PRIOR APPLICATION NUMBER: US/09/967,772

PRIOR FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: US 09/459,956

PRIOR FILING DATE: 1999-12-13

PRIOR APPLICATION NUMBER: US 08/765,860

PRIOR FILING DATE: 1996-12-19

PRIOR APPLICATION NUMBER: PCT/ US96/09652

PRIOR FILING DATE: 1996-06-06

PRIOR APPLICATION NUMBER: US 08/481,977

; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-10-335-517-6

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-335-517-6 (1-678)

QY 1 MetArgSerSerLySAsnValIleLysGluPheMetArgPheLySValArgMetGluGly 20  
1 ATGAGGCTCTTCCAAGAAATGTTATCAAGAGATTCAAGGTTTAAGTTTCCGATGGAAAGCA 60  
DB 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
61 ACGGTCAATGGGCACGAGTTTGAATAAGAGCGAAGAGAGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGLYGLYProLeuProPheAlaTrpAspIle 60  
121 CACAATAACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTGGCATTTGCTTGGGATATT 180  
DB 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
181 TTGTCAACCACAATTTCAAGTATGGAAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
241 GACTATATAAAAGCTGTATTCCTGAAGGATTAAATGGGAAAGGTCATGAACCTTTGAA 300  
DB 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
301 GACGGTGGCGTCTTACTGTAAACCCAGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
361 AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGAGACCTGTTATGCAAAAGAAAGACA 420  
DB 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLYGLU 160  
421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAGAGAGAG 480  
QY 161 IleHisValAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuThr 180  
481 ATTCATTAAGGCTCTGAAGCTGAAAGACGGTGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
DB 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
541 TACATGGCAAGAACCTGTGACAGTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
601 ATAACAAGCCACAAGAACTATACAAATCGTTGAGACAGTATGAAGAACCAGAGGACGC 660  
DB 221 HisHisLeuPheLeu 225  
661 CACCATCTGTTCTT 675

RESULT 8  
US-10-334-288-6  
Sequence 6, Application US/10334288  
Publication No. US20040002123A1

GENERAL INFORMATION:

APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

APPLICANT: TSJEN, Roger







QY 161 ILEHISLYSALALEULYSLEULYSASPGLYGLYHISTYRLEUVALGLUPHELYSSEITIE 180  
DB 481 ATTCAITAAAGGCTCTGAAGCTGMAAGACGGGTGTCATTACCTAGTTGAATTCAAAGATATT 540  
QY 181 TYMECALALYSLSYSPROVALGINLEUPROGLYTYRTYTYRVALASPSERLYSLEUASP 200  
DB 541 TACATGGCAAGAACCTGTGTGCACTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 ILETHERSERHISASNGLUASPTYRTHRILEVALGLUGLINTYRGUARGTARGLUGLYARG 220  
DB 601 ATAAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGAGGACGC 660  
QY 221 HISHISLEUPHELEU 225  
DB 661 CACCATCTGTTCCTT 675

RESULT 10  
US-10-656-029-21  
; Sequence 21, Application US/10656029  
; Publication No. US2005000367A1  
; GENERAL INFORMATION:  
; APPLICANT: VERTEX PHARMACEUTICALS INC.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF  
; TITLE OF INVENTION: SCREENING ASSAYS  
; FILE REFERENCE: VPI/02-143WO2  
; CURRENT APPLICATION NUMBER: US/10/656,029  
; CURRENT FILING DATE: 2003-09-05  
; PRIOR APPLICATION NUMBER: 60/408,297  
; PRIOR FILING DATE: 2002-09-05  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 21  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: DISCOSOMA SP.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(678)  
; OTHER INFORMATION: fluorescent protein  
; US-10-656-029-21

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-656-029-21 (1-678)

QY 1 METARGSERSELYSASNVALILEYSGIUPHEMETARGPHELYSVALARGMETGLUGLY 20  
DB 1 ATGAGGTCTTCCAGAATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCCGATGGAAGGA 60  
QY 21 THRVALASNGLYHISGLUPHEGLUILEGUGLYGLUGLYARGPROTYRGLUGLY 40  
DB 61 ACGGTCAATGGGCAAGAGTTTGAATAAGAGCGAAGAGGAGGAGCCATACGAAGGC 120  
QY 41 HISASNTHRVALYSLEULYSVALTHRLYSGLYGLYPROLEUPROPHALATRPASPILE 60  
DB 121 CACAATACCGTAACCTTAAGGTAACCAAGGGGGAACCTTGCCATTGGCTGGGATATT 180  
QY 61 LEUSERPROGLINPHEGLINTYRGLYSERLYSVALTYRVALLYSHISPROALAASPILEPRO 80  
DB 181 TTGTCACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCAACCTTGCCACATACCA 240  
QY 81 ASPTYRLYSALYSLEUSERPHEPROGLUGLYPHELYSTRPGLUARGVALMETASNPHGLU 100  
DB 241 GACTATATAAAGCTGTCAATTCTCGAAGGATTTAATGGGAAAGGGTCAATGAACCTTTGAA 300  
QY 101 ASPGLYGLYVALVALTHRVALTHRGLINASPSESRLEUGLINASPGLYCYSPHEILETYR 120

DB 301 GACGGTGGCGCTGTTACTGTAAACCCAGGATTCAGTTTGACGAGATGGCTGTTTCAITCTAC 360  
QY 121 LYSVALLYSPHEILEGLYVALASNPHEPROSERASPGLYPROVALMETGLNLYSSTR 140  
DB 361 AAGGTCAAGTTCAATGGCGGAACCTTCTCCGATGGACCTGTATGCAAAAGAGACA 420  
QY 141 METGLYTRPGLUALASERTHRGLUARGLEUTYRPROARGASPGLYVALLEULYSGLYGLU 160  
DB 421 ATGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTTGAAGAGAG 480  
QY 161 ILEHISLYSALALEULYSLEULYSASPGLYGLYHISTYRLEUVALGLUPHELYSSEITIE 180  
DB 481 ATTCAITAAAGCTCTGAAGCTGAAGAAGACGGGTGTCATTACCTAGTTGAATTCAAAGATATT 540  
QY 181 TYMECALALYSLSYSPROVALGINLEUPROGLYTYRTYTYRVALASPSERLYSLEUASP 200  
DB 541 TACATGGCAAGAACCTGTGTGCACTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 ILETHERSERHISASNGLUASPTYRTHRILEVALGLUGLINTYRGUARGTARGLUGLYARG 220  
DB 601 ATAAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGAGGACGC 660  
QY 221 HISHISLEUPHELEU 225  
DB 661 CACCATCTGTTCCTT 675

RESULT 11  
US-10-505-486-27  
; Sequence 27, Application US/10505486  
; Publication No. US20050118639A1  
; GENERAL INFORMATION:  
; APPLICANT: Takeda Chemical Industries, Ltd.  
; TITLE OF INVENTION: Determination of a ligand  
; FILE REFERENCE: P03-0006PCT  
; CURRENT APPLICATION NUMBER: US/10/505,486  
; CURRENT FILING DATE: 2004-08-20  
; PRIOR APPLICATION NUMBER: JP 2002-45728  
; PRIOR FILING DATE: 2002-02-22  
; PRIOR APPLICATION NUMBER: JP 2002-213949  
; PRIOR FILING DATE: 2002-07-23  
; PRIOR APPLICATION NUMBER: JP 2002-298237  
; PRIOR FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 233  
; SEQ ID NO 27  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: DISCOSOMA SP.  
; US-10-505-486-27

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-505-486-27 (1-678)

QY 1 METARGSERSELYSASNVALILEYSGIUPHEMETARGPHELYSVALARGMETGLUGLY 20  
DB 1 ATGAGGTCTTCCAGAATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCCGATGGAAGGA 60  
QY 21 THRVALASNGLYHISGLUPHEGLUILEGUGLYGLUGLYARGPROTYRGLUGLY 40  
DB 61 ACGGTCAATGGGCAAGAGTTTGAATAAGAGCGAAGAGGAGGAGCCATACGAAGGC 120  
QY 41 HISASNTHRVALYSLEULYSVALTHRLYSGLYGLYPROLEUPROPHALATRPASPILE 60  
DB 121 CACAATACCGTAACCTTAAGGTAACCAAGGGGGAACCTTGCCATTGGGATATT 180  
QY 61 LEUSERPROGLINPHEGLINTYRGLYSERLYSVALTYRVALLYSHISPROALAASPILEPRO 80



Db 181 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPhelGlu 100

Db 241 GACTATAAAAAAGCTGTCAATTTCTGAAGGATTAAATGGGAAAGGTCATGAACCTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 301 GACGGTGGCGTCTTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 361 AAGTCAAGTTCATTGGCGTAACCTTCCGATGGAACCTGTTATGCAAAAGAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 421 ATGGGCTGGGAAGCCAGCACTGACGCTTGTATCCTCGTGAATGGCGTGTGAAGAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 541 TACATGGCAAAAGAGCCTGTGCAGCTACAGGGTACTATAGTTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAATGAAGAACCAGAGGAGCG 660

QY 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTCTCTT 675

RESULT 12

US-10-844-064A-1

/ Sequence 1, Application US/10844064A

/ Publication No. US20050149994A1

/ GENERAL INFORMATION:

/ APPLICANT: Bevis, Brooke

/ APPLICANT: Glick, Benjamin

/ TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S

/ FILE REFERENCE: 092234-9006

/ CURRENT APPLICATION NUMBER: US/10/844, 064A

/ PRIOR FILING DATE: 2004-05-11

/ PRIOR APPLICATION NUMBER: PCT/US02/40539

/ PRIOR FILING DATE: 2002-12-18

/ PRIOR APPLICATION NUMBER: US 60/341, 723

/ PRIOR FILING DATE: 2001-12-19

/ NUMBER OF SEQ ID NOS: 3

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 1

/ LENGTH: 678

/ TYPE: DNA

/ ORGANISM: Discosoma

US-10-844-064A-1

Alignment Scores:

Pred. No.: 6.83e-143 Length: 678

Score: 1214.00 Matches: 225

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-844-064A-1 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyGly 20

Db 1 ATGAGGTCTTCCAAAGAAATGTTATCAAGAGATTCAATGAGTTTAAGGTTCCGATGGAAGGA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyArgProTyrGlyGly 40

Db 61 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGGCCAAGAGAGGGGAGGCCCATACGAAGGC 120

QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

Db 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGTGCTGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 181 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPhelGlu 100

Db 241 GACTATAAAAAAGCTGTCAATTTCTGAAGGATTAAATGGGAAGGTCATGAACCTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 301 GACGGTGGCGTCTTACTGTAAACCAAGATTCAGATTTCAGAGATGGCTGTTTCATCTAC 360

QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 361 AAGTCAAGTTCAATTGGCGTGAACCTTCTCCGATGGAACCTGTTATGCAAAAGAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTAAGAAGAG 480

QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTAATACTATGTTGACTCCAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220

Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAATGAAGAACCAGAGGAGCG 660

QY 221 HisHisLeuPheLeu 225

Db 661 CACCATCTGTCTCTT 675

RESULT 13

US-10-931-304-2

/ Sequence 2, Application US/10931304

/ Publication No. US20050196768A1

/ GENERAL INFORMATION:

/ APPLICANT: Tsien, Roger

/ APPLICANT: Campbell, Robert

/ APPLICANT: Geofrey

/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

/ FILE REFERENCE: 39754-0831CP2CP3

/ CURRENT APPLICATION NUMBER: US/10/931, 304

/ PRIOR FILING DATE: 2004-08-30

/ PRIOR APPLICATION NUMBER: 10/209, 208

/ PRIOR FILING DATE: 2002-07-29

/ PRIOR APPLICATION NUMBER: 10/121, 258

/ PRIOR FILING DATE: 2002-04-10

/ PRIOR APPLICATION NUMBER: 09/866, 538

/ PRIOR FILING DATE: 2001-05-24

/ PRIOR APPLICATION NUMBER: 09/794, 308

/ PRIOR FILING DATE: 2001-02-26

/ NUMBER OF SEQ ID NOS: 110

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 2

/ LENGTH: 678

/ TYPE: DNA

/ ORGANISM: Discosoma sp.

/ FEATURE:

/ NAME/KEY: misc\_feature

/ LOCATION: (1)...(678)



OTHER INFORMATION: wild-type DsRed  
US-10-931-304-2

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-2 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGCTCTCCAGAATGTATCAAGAGGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGAAGGAGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTGGCCATTGCGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAGCTGTCTATTCTCGAAGGATTTAAATGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCCAGAGTTCCAGTTGCAGAGTGGCTGTTTCATCTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTTCATTGGCGTGAACTTCTTCGATGGACCTGTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGATGGCGGTTTGAAAGAGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
DB 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCATTACCTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAACCTGTGCAGCTACCGGTACTACTATGTGACTCCAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAAGAGACTATACATCGTTGAGCAGATGAAAGAACCGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCTT 675

RESULT 14

US-09-999-745-66  
Sequence 66, Application US/09999745  
Patent No. US20020157120A1  
GENERAL INFORMATION:  
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: Telen, Roger Y.  
APPLICANT: Baird, Geoffrey  
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
FILE REFERENCE: REGEN1470-1  
CURRENT APPLICATION NUMBER: US/09/999,745  
CURRENT FILING DATE: 2001-10-23

PRIOR APPLICATION NUMBER: 09/316,920  
PRIOR FILING DATE: 1999-05-21  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 66  
LENGTH: 859  
TYPE: DNA  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54) .. (731)  
US-09-999-745-66

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-999-745-66 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 54 ATGAGGCTCTCCAGAATGTATCAAGAGGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCACGAGTTTGAAATAGAAGCGAAGGAGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 174 CACAAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTGGCCATTGCGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCCGACATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 294 GACTATAAAGCTGTCTATTCTCGAAGGATTTAAATGGAAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGTGGCGTCTTACTGTAAACCCAGAGTTCCAGTTGCAGAGTGGCTGTTTCATCTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 414 AAGGTCAAGTTTCATTGGCGTGAACTTCTTCGATGGAACCTGTATGCAAAAGAGACA 473  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGATGGCGGTTTGAAAGAGAG 533  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
DB 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCATTACCTAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 594 TACATGGCAAGAACCTGTGCAGCTACCGGTACTACTATGTGACTCCAAACTGGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATAACAAGCCACAAGAGACTATACATCGTTGAGCAGATGAAAGAACCGAGGAGCGC 713  
QY 221 HisHisLeuPheLeu 225  
DB 714 CACCATCTGTTCCTT 728

RESULT 15

US-09-866-538-11



/ Sequence 11, Application US/09866538  
/ Publication No. US20030032088A1  
/ GENERAL INFORMATION:  
/ APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
/ APPLICANT: TSIEN, Roger  
/ APPLICANT: Campbell, Robert  
/ TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
/ FILE REFERENCE: REGEN1530-2  
/ CURRENT APPLICATION NUMBER: US/09/866,538  
/ CURRENT FILING DATE: 2001-05-24  
/ NUMBER OF SEQ ID NOS: 29  
/ SOFTWARE: PatentIn version 3.0  
/ SEQ ID NO 11  
/ LENGTH: 859  
/ TYPE: DNA  
/ ORGANISM: Discosoma sp.  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (54)..(731)  
US-09-866-538-11

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-866-538-11 (1-859)

QY 1 MetArgSerSerLyAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGGTCTTCCAGAATGTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGCAGCAGATTGAAATAGAAAGGCCAAGAGAGAGGGGACCATTTGGGATATT 173  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATACCGTAAGCTTAAGTTAACCAAGGGGGGACCTTTGCCATTGTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 234 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGCAAGCACCCGCGACATACCA 293  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAGCTGTCAATTCCTGAAGATTAAATGGAAAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCTGTAATCTGAACCCAGGATTCAGTTTGCAAGATGGCTGTTCACTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTCATTTGGCGTGAACCTTCTCCGATGACCTGTTATGCMAAAGAAACA 473  
QY 141 MetGlyTyrGluAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGACCGTTTGTATCCTCGTATGCGGTGTGAAGAGAG 533  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTGTGAAGCTGAAGAAGCGGTGTCATTAAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAAGCCTGTGACGCTACCAAGGGTACTATATGTTGACTCCAAGCTGGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 220

Db 654 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAGAAGCGAGGAGCGC 713  
QY 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCCTT 728

Search completed: January 12, 2006, 12:40:06  
Job time : 797 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 10:16:32 ; Search time 383 Seconds  
(without alignments)  
475.466 Million cell updates/sec

Title: US-10-006-922a-12  
Perfect score: 1214  
Sequence: 1 MRSKNIKEFMRFKVRMEG.....EDYTIQEYRTGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_10012006\_162410\_19008/app\_query.fasta\_1.391  
-DB=Published\_Applications\_NA\_New -QFMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10006922\_@CGN\_1\_1\_184\_@runat\_10012006\_162410\_19008  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA\_New:\*  
1: /cgn2\_6/pcodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/pcodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/pcodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/pcodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/pcodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/pcodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/pcodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/pcodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/pcodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/pcodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	US-10-209-208-2	Sequence 2, Appli
2	1214	100.0	678	US-11-218-880-2	Sequence 2, Appli
3	1210	99.7	681	US-10-209-208-3	Sequence 3, Appli
4	1210	99.7	681	US-10-209-208-23	Sequence 3, Appli
5	1210	99.7	681	US-11-218-880-3	Sequence 23, Appli
6	1210	99.7	681	US-11-218-880-23	Sequence 23, Appli
7	1186	97.7	6706	US-10-655-872-4	Sequence 4, Appli
8	1186	97.7	7927	US-10-655-872-7	Sequence 7, Appli

9	1160	95.6	678	6	US-10-209-208-5	Sequence 5, Appli
10	1160	95.6	678	7	US-11-218-880-5	Sequence 5, Appli
11	1121	92.3	681	6	US-10-209-208-7	Sequence 7, Appli
12	1121	92.3	681	7	US-11-218-880-7	Sequence 7, Appli
13	1021	84.1	678	6	US-10-209-208-9	Sequence 9, Appli
14	1021	84.1	678	7	US-11-218-880-9	Sequence 9, Appli
15	1012	83.4	678	6	US-10-209-208-80	Sequence 80, Appli
16	486.5	40.1	1079	7	US-11-179-411-15	Sequence 15, Appli
17	486.5	40.1	1079	7	US-11-175-766-15	Sequence 15, Appli
18	474.5	39.1	1104	7	US-11-179-411-30	Sequence 30, Appli
19	474.5	39.1	1104	7	US-11-175-766-30	Sequence 30, Appli
20	474.5	39.1	1279	7	US-11-179-411-31	Sequence 31, Appli
21	474.5	39.1	1279	7	US-11-175-766-31	Sequence 31, Appli
22	468.5	38.6	6679	6	US-10-521-768-1	Sequence 1, Appli
23	468.5	38.6	8251	6	US-10-521-768-2	Sequence 2, Appli
24	468.5	38.6	10369	6	US-10-521-768-3	Sequence 3, Appli
25	468.5	38.6	13535	6	US-10-521-768-4	Sequence 4, Appli
26	422.5	34.8	860	7	US-11-179-411-24	Sequence 24, Appli
27	422.5	34.8	860	7	US-11-175-766-24	Sequence 24, Appli
28	422.5	34.8	864	7	US-11-179-411-23	Sequence 23, Appli
29	422.5	34.8	864	7	US-11-179-411-26	Sequence 26, Appli
30	422.5	34.8	864	7	US-11-175-766-23	Sequence 23, Appli
31	422.5	34.8	864	7	US-11-175-766-26	Sequence 26, Appli
32	422.5	34.8	873	7	US-11-179-411-25	Sequence 25, Appli
33	422.5	34.8	873	7	US-11-175-766-25	Sequence 25, Appli
34	219.5	18.1	10122	6	US-10-161-408-8	Sequence 8, Appli
35	219.5	18.1	10549	6	US-10-161-408-3	Sequence 3, Appli
36	219.5	18.1	12592	6	US-10-161-408-6	Sequence 6, Appli
37	210.5	17.3	716	6	US-10-209-208-21	Sequence 21, Appli
38	210.5	17.3	716	7	US-11-218-880-21	Sequence 21, Appli
39	210.5	17.3	717	7	US-11-089-551A-12	Sequence 12, Appli
40	210.5	17.3	720	7	US-11-089-551A-39	Sequence 39, Appli
41	210.5	17.3	7239	7	US-11-089-551A-45	Sequence 45, Appli
42	209.5	17.3	4773	7	US-11-214-613-32	Sequence 32, Appli
43	209.5	17.3	5164	7	US-11-214-613-36	Sequence 36, Appli
44	204.5	16.8	1539	7	US-11-032-236-5	Sequence 5, Appli
45	204.5	16.8	3660	7	US-11-032-236-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-10-209-208-2  
; Sequence 2, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(678)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-2

Alignment Scores:



Pred. No.:	2.23e-133	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x US-10-209-208-2 (1-678)

QY	1	MetArgSerSerLySAsnValIleIysGluPheMetArgPheLySValArgMetGlu	20
Db	1	ATGAGGCTTCCAAAGATGTTATCAAGAGCTTCATGAGGTTAAAGGTTCCATGGAAGCA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCGAAGAGAGGGAGGCCATACGAAGC	120
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
Db	121	CACAATACCGTAAGCTTAAGGTAACCAAGGGGGACCTTGCCATTGCTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro	80
Db	181	TTGTCAACCAATTTCACTATGGAAGCAAGTATATGTCAAGCAACCTGCCGACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu	100
Db	241	GACTATATAAAGCTGTCACTTCTCGAAGATTAAATGGGAAAGGTCATGAACTTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGTGTACTGTAAACCAAGATTCAGTTTCAGAGATGGCTGTTCAATCTAC	360
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr	140
Db	361	AAGGTCAAGTTCACTGGCGTGAACCTTCTCCGATGGAACCTGTATGCAAGAAGACA	420
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160
Db	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGTGTGAAGAGAGAG	480
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyHisTyrLeuValGluPheLySLeu	180
Db	481	ATTCAATAAGGCTCTGAAGCTGAAGACGCGTGCTCATTAACCTAGTTGAATTCAAAATATT	540
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrValAspSerLySLeuAsp	200
Db	541	TACATGGCAAGAAGCCTGTGACGTAACCAAGGGTACTACTATGTGACTCCTGAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACAAGACTATACATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCATCTGTTCCCTT 675	
RESULT 2			
US-11-218-880-2			
; Sequence 2, Application US/11218880			
; Publication No. US20060003420A1			
GENERAL INFORMATION:			
; APPLICANT: Tsien, Roger			
; APPLICANT: Campbell, Robert			
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT			
; FILE REFERENCE: UC083.1CP2CPI			
; CURRENT APPLICATION NUMBER: US/11/218,880			
; CURRENT FILING DATE: 2005-09-01			
; PRIOR APPLICATION NUMBER: US/10/121,258			
; PRIOR FILING DATE: 2002-04-10			
; PRIOR APPLICATION NUMBER: 09/794,308			
; PRIOR FILING DATE: 2001-02-26			

```

; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (678)
; OTHER INFORMATION: wild-type Dated
US-11-218-880-2

```

### Alignment Scores:

Pred. No.:	2.23e-133	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-10-006-922A-12 (1-225) x US-11-218-880-2 (1-678)

Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGAGGCTTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAAGTTCCATGGAAGGA	60
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACGGTCAATGGGCACGAGTTTGAAATGAAGGCGAAGAGAGAGGGGCCATTACGAAGGC	120
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACATACCGTAAAGCTTAAGGTAACCAAGGGGGGACCTTGGCATTTGCTGGGATATT	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	TTGTCAACCACAAATTTCAGTATGGAAGCAAGGATATATGTCACAGCACCTGCCGACATACCA	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTATAAAAAGCTGTCAATTCCTGGAAGATTAAATGGGAAAGGTCATGAACCTTGAA	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGCTGTACTGTAAACCAAGGATTCACAGTTTGCAGGATGGCTGTTTCACTTAC	360
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTCAAGTTCAATGGCGGTGAACCTTCTCCGATGACCTGTTATGCCAAAAGAAAGACA	420
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTATGCGCTGTGAAAAGAGAG	480
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTTACTAGTTGAATTCAAAAAGTATT	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCAAAGAGCGCTGTGACGCTACGAGGTACTACTATGTTGACTCCAAACTGGAT	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACACGAAGACTATATCAATCGTTGAGCAGTATGAAGAACCAGAGGACGC	660
Qy	221	HisHisLeuPheLeu	225
Db	661	CACCATCTGTTCTT	675



US-10-209-208-3  
; Sequence 3, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-10-209-208-3

Alignment Scores:  
Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-3 (1-681)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB ::  
4 GTGCGCTCCTCCAAAGACGTCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB ::  
64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB ::  
124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGGGCCCCCTGCGCTTGGCGATCGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB ::  
184 CTGTCCCCCCCAGTTCCAGTACGGCTCCAAAGTGTACGTGAAGCACCCCCGCAATCCCC 243  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB ::  
244 GACTACAAGAAAGTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGGTGATGAATTGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
DB ::  
304 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCAGAGCGGCTGCTCATCTAC 363  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB ::  
364 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAAAGACC 423  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB ::  
424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGAGCGGCGTGTGAAGGGCGAG 483  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB ::

DB 484 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 543  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrValAspSerLySLeuAsp 200  
DB ::  
544 TACATGGCCAAGAGCCGTGACAGTGCCTCCGCTACTACTACGTGACTCCAGCTGGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB ::  
604 ATCACTCCCAACAGAGACTACACCATCTGTGAGAGTAGAGGCGCACCGAGGGCCGC 663  
QY 221 HisHisLeuPheLeu 225  
DB 664 CACCACCTGTCTCTG 678

RESULT 4

US-10-209-208-23  
; Sequence 23, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-10-209-208-23

Alignment Scores:  
Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-23 (1-681)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB ::  
4 GTGCGCTCCTCCAAAGACGTCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB ::  
64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB ::  
124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB ::  
184 CTGTCCCCCAGTTCAGTACGGCTCCAAAGTGTACGTGAAGCACCCCCGCAATCCCC 243  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB ::  
244 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTGAG 303



QY 101 AaPgLyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||  
Db 304 GACGGCGCGCTGGTGACCGGTGACCCAGGACTCTCCCTGCGAGGACGGCTGCTTCACTTAC 363  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||  
Db 364 AAGGTGAAGTTCAATCGCGGTGAACTTCCCTCCGACGGCCCGTAATGCAGAGAGAAC 423  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyLys 160  
|||  
Db 424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGGTGTAAGGCGAG 483  
QY 161 LLeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTGGTGAAGTCAAGTCCATC 543  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||  
Db 544 TACATGGCCCAAGAGCCCGGTGACGTGCGGCTACTACTACTGGACTCCAAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||  
Db 604 ATCACTCCCAACAAGGAGACTACCATCTGTGAGCAGTACGAGCGCACCGGAGGGCGGC 663  
QY 221 HisHisLeuPheLeu 225  
|||  
Db 664 CACCACCTGTTCTTG 678

RESULT 5  
US-11-218-880-3  
; Sequence 3, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-11-218-880-3

Alignment Scores:  
Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-3 (1-681)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyLys 20  
:::|||||  
Db 4 GTGCGCTCCTCCAAAGAACGTCAACAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 64 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
|||  
Db 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTCCCTTGGGAGATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
|||  
Db 184 CTGTCCCGCCAGTTCCAGTACGAGCTCCAAAGGTGTAAGTGAAGCACCCCGCCGATCCCC 243  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
|||  
Db 244 GACTACAAAGAGCTGTCTTCCCGGAGGCTTCAAGTGGAGCGCGTGAATGAATTGAG 303  
QY 101 AaPgLyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||  
Db 304 GACGGCGCGGTGGTGACCGTGACCCAGACTCTCCCTGACGAGACGGCTGCTCATCTAC 363  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
|||  
Db 364 AAGGTGAAGTTCAATCGCGGTGAACTTCCCTCCGACGGCCCGTAATGCAGAGAGAAC 423  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyLys 160  
|||  
Db 424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGGTGTAAGGCGAG 483  
QY 161 LLeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTGGTGAAGTCAAGTCCATC 543  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||  
Db 544 TACATGGCCCAAGAGCCCGGTGACGTGCGGCTACTACTACTGAGTCCAAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||  
Db 604 ATCACTCCCAACAAGGAGACTACCATCTGTGAGCAGTACGAGCGCACCGGAGGGCGGC 663  
QY 221 HisHisLeuPheLeu 225  
|||  
Db 664 CACCACCTGTTCTTG 678

RESULT 6  
US-11-218-880-23  
; Sequence 23, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-11-218-880-23

Alignment Scores:



Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-23 (1-681).

```
QY      1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20
DB      4 GTGCGCTCCTCCAAGACGTCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 63
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB      64 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGGC 123
QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
DB      124 CACAAACCCGTGAAGCTGAAGGTGACCAAGGGGGGGCCCTGCGCTTGGGACATC 183
QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
DB      184 CTGTCCCCCAAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCAACCCCGGACATCCCC 243
QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB      244 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 303
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      304 GACGGCGCGGTGTGACCGGTGACCCAGACTCTCCCTGCAGAGCGGCTGCTTCATCTAC 363
QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
DB      364 AAGGTGAAGTTCATCGCGGTGAACTTCCCTCCGACGGCCCGTAATGCAGAAAGAACCC 423
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
DB      424 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGGTGAAAGGGCGAG 483
QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
DB      484 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGGCACTACCTGTGTGAGTTCAAGTCCATC 543
QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
DB      544 TACATGGCCAAAGAGCCCGTGCAGCTGCCGGCTACTACTACGTGACTCCAAGCTGAC 603
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      604 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGGC 663
QY      221 HisHisLeuPheLeu 225
DB      664 CACCACCTGTTCTTG 678
```

RESULT 7

US-10-655-872-4  
; Sequence 4, Application US/10655872  
; Publication No. US20050251872A1  
; GENERAL INFORMATION:  
; APPLICANT: Bear, et al.  
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
; FILE REFERENCE: 0492611-0512  
; CURRENT APPLICATION NUMBER: US/10/655,872  
; CURRENT FILING DATE: 2003-09-05  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 6706  
; TYPE: DNA  
; ORGANISM: Artificial

FEATURE:  
; OTHER INFORMATION: lentiviral vector sequence  
; US-10-655-872-4

Alignment Scores:  
Pred. No.: 1e-128 Length: 6706  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-655-872-4 (1-6706)

```
QY      1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20
DB      2778 ATGGCTCCTCCGAGAACGTCAACCGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 2837
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB      2838 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGGC 2897
QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
DB      2898 CACAAACCCGTGAAGCTGAAGGTGACCAAGGGGGGGCCCTGCGCTTGGGACATC 2957
QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
DB      2958 CTGTCCCCCAAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCAACCCCGGACATCCCC 3017
QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB      3018 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 3077
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      3078 GACGGCGCGGTGGCGAGCGGTGACCAAGACTCCTCCCTGCAGAGCGGCTGCTTCATCTAC 3137
QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
DB      3138 AAGGTGAAGTTCATCGCGGTGAACTTCCCTCCGACGGCCCGGTGATGCAGAAAGAACCC 3197
QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
DB      3198 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGGTGAAAGGGCGAG 3257
QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
DB      3258 ACCCAAGAGGCCCTGAAGCTGAAGAGACGGCGGCACTACCTGTGTGAGTTCAAGTCCATC 3317
QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
DB      3318 TACATGGCCAAAGAGCCCGTGCAGCTGCCGGCTACTACTACGTGAGCAGCAAGCTGAC 3377
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      3378 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGGC 3437
QY      221 HisHisLeuPheLeu 225
DB      3438 CACCACCTGTTCTTG 3452
```

RESULT 8

US-10-655-872-7  
; Sequence 7, Application US/10655872  
; Publication No. US20050251872A1  
; GENERAL INFORMATION:  
; APPLICANT: Bear, et al.  
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
; FILE REFERENCE: 0492611-0512  
; CURRENT APPLICATION NUMBER: US/10/655,872  
; CURRENT FILING DATE: 2003-09-05  
; NUMBER OF SEQ ID NOS: 39



```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 7927
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: lentiviral vector sequence
US-10-655-872-7

Alignment Scores:
Pred. No.:      1.26e-128      Length:      7927
Score:          1186.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%      Mismatches:  4
Query Match:    97.69%          Indels:       0
DB:             6              Gaps:         0

US-10-006-922a-12 (1-225) x US-10-655-872-7 (1-7927)

QY      1 MetArgSerSerLybAbnValIlleYsGluPhMeTArgPhelybValArgMetGluGly 20
Db      3999 ATGGCCTCCTCCGAGAACGTCATCACCGAGTTCATGCCCTTCAAGGTGCGCATGGAGGGC 4058

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      4059 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGAGGGCCCCCTACGAGGGC 4118

QY      41 HisAbnThrValYsLeuYsValThrLYsGlyGlyProLeuProPheAlaTrpAspIle 60
Db      4119 CACAAACACCGTGAAGCTGAAGGTGAACCAAGGGCGGGCCCCCTGCCCTTGCGGACATC 4178

QY      61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLYsHisProAlaAspIlePro 80
Db      4179 CTGTCCCCCCAGTTCAGATCGAGCTCCAAGGTGTAAGTAAGCAACCCCGCAGCATCCCC 4238

QY      81 AspTyrLYsLYsLeuSerPheProGluGlyPheLYsTrpGluArgValMetAsnPhelGlu 100
Db      4239 GACTACAAGAGCTGTCTCTTCCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACTTCCAG 4298

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
Db      4299 GACGGCGCGCTGGCGACCGCTGACCCAGGACTCTCTCTGCAAGACGGCTGCTTCACTTAC 4358

QY      121 LybValLYsPheIleGlyValAsnPheProSerAspGlyProValMetGlnLYsLYsThr 140
Db      4359 AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAAGACC 4418

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuYsGlyGlu 160
Db      4419 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGCAGCGCGTGTGAAGGCGAG 4478

QY      161 IleHisLYsAlaLeuLYsLeuLYsAspGlyGlyHisTyrLeuValGluPheLYsSerIle 180
Db      4479 ACCCAACAAGGCCCTGAAGCTGAAGAGCGGCGCCACTACCTGGTGAAGTTCAGTCCATC 4538

QY      181 TyrMetAlaLYsLYsProValGlnLeuProGlyTyrTyrTyrValAspSerLYsLeuAsp 200
Db      4539 TACATGGCCCAAGAGCCCGCTGCAGCTGCCCGGCTACTACTACGTGGAACGCCAAGCTGCAC 4598

QY      201 IleThrSerHisAbnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      4599 ATCACTCCCAACAAGAGGACTACCATCTGTGAGAGTACGAGCGCACCGAGGGCGCGC 4658

QY      221 HisHisLeuPheLeu 225
Db      4659 CACCACTGTTCCTG 4673

RESULT 9
US-10-209-208-5
; Sequence 5, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
```

```

; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey, Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "T1"
US-10-209-208-5

Alignment Scores:
Pred. No.:      5.3e-127      Length:      678
Score:          1160.00        Matches:      216
Percent Similarity: 97.33%      Conservative: 3
Best Local Similarity: 96.00%      Mismatches:  6
Query Match:    95.55%          Indels:       0
DB:             6              Gaps:         0

US-10-006-922a-12 (1-225) x US-10-209-208-5 (1-678)

QY      1 MetArgSerSerLybAbnValIlleYsGluPhMeTArgPhelybValArgMetGluGly 20
Db      1 ATGGCCTCCTCCGAGAACGTCATCAAGAGTTCATGCCCTTCAAGGTGCGCATGGAGGGC 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      61 TCCGTGAACGGCCACGAGTTCAGATCGAGGCGAGGGCGCGCCCTTACGAGGGC 120

QY      41 HisAbnThrValLYsLeuLYsValThrLYsGlyGlyProLeuProPheAlaTrpAspIle 60
Db      121 ACCCAGACCGCGAAGCTGAAGTGACCAAGGGCGGGCCCCCTGCCCTGCGGAGACATC 180

QY      61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLYsHisProAlaAspIlePro 80
Db      181 CTGTCCCCCCAGTTCAGTACGGCTCCAAAGGTGTAAGTGAAGCAACCCCGCAGCATCCCC 240

QY      81 AspTyrLYsLYsLeuSerPheProGluGlyPheLYsTrpGluArgValMetAsnPhelGlu 100
Db      241 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACCTTCGAG 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
Db      301 GACGGCGCGCTGGTGAACCGTGACCAAGACTCTCTCCCTGACGAGCGGCTCTTCATCTAC 360

QY      121 LybValLYsPheIleGlyValAsnPheProSerAspGlyProValMetGlnLYsLYsThr 140
Db      361 AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGGCCCCGTAAATGCAAGAAAGACT 420

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuYsGlyGlu 160
Db      421 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGGTGTAAGGGCGAG 480

QY      161 IleHisLYsAlaLeuLYsLeuLYsAspGlyGlyHisTyrLeuValGluPheLYsSerIle 180
Db      481 ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGCCACTACCTGTGAGTTCAGATCCATC 540

QY      181 TyrMetAlaLYsLYsProValGlnLeuProGlyTyrTyrTyrValAspSerLYsLeuAsp 200
Db      541 TACATGGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAGCTGGAC 600
```



QY 201 ILeThSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluYArg 220  
Db 601 ATCACCCTCCCAACAGAGACTACCATCTGTGAGACGAGTACGAGCGCGGAGGCGCCG 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTTG 675

RESULT 10

US-11-218-880-5  
; Sequence 5, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "T1"  
US-11-218-880-5

Alignment Scores:  
Pred. No.: 5.3e-127 Length: 678  
Score: 1160.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.55% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-5 (1-678)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
Db 1 ATGGCCTCCTCCGAGGACGTTCATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 TCCGTGAACGCGCACGAGTTGAGATCGAGGCGGAGGCGGCGCCCTACGAGGGC 120  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGCGCGCCCTGCGCTTCCGCGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCAGTACGCGCTCCAAGGTGATGAGCAACCCGCGGACATCCCC 240  
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAAGCTGTCCTTCCCGGAGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 300  
QY 101 AsnGlyGlyValIleValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGTGACCGTGAACCCAGGACTCCTCCCTGCGAGGACGGCTCCTTCACTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140

Db 361 AAGGTGAAGTTTCATCGCGGTGAATTCCCTCCGACGCGCCCGTAATGCAAGAGACT 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIle 160  
Db 421 ATGGGTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 480  
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyGlyHisIstYrLeuValGluPheIysSerIle 180  
Db 481 ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGCACCTACTGCTGAGTTCAAGTTCATC 540  
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
Db 541 TACATGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGAAGTCCAGCTGAG 600  
QY 201 ILeThSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluYArg 220  
Db 601 ATCACCCTCCCAACAGAGACTACCATCTGTGAGACGAGTACGAGCGCGGAGGCGCCG 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTTG 675

RESULT 11

US-10-209-208-7  
; Sequence 7, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "dimer2"  
US-10-209-208-7

Alignment Scores:  
Pred. No.: 2.15e-122 Length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservative: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-7 (1-681)

QY 3 SerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGlyThrVal 22  
Db 10 TCCTCCGAGGACGTTCATAAAGATTGCGCTTCAAGGTGCGCATGAGGCGTCCGTG 69  
QY 23 AsnGlyHisGluPheGluIleGluGlyGlyGlyArgProTyrGluGlyHisAsn 42  
Db 70 AACGGCCACGAGTTGAGATCGAGGCGGAGGCGGCGCCCTACGAGGACCAACCAG 129  
QY 43 ThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62



Db 130 ACCGCCAAGCTGAAGTGACGAAGGGCGGCCCTGCGCTTCGCTGGAGACATCTGTCC 189  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
190 CCCAGTTCACAGTACGGCTCCAAAGCGGTACGTGAAGCACCCCGCGACATCCCGGACTAC 249  
QY 83 LysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGluAspGly 102  
250 AAGAAGCTGTCTCTCCCGGAGGCTTCAAGTGGAGCGCGGTGATGAATCTCGAGGACGGC 309  
Db 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
310 GCGGTGGTGACCGTGACCCAGGACTCTCCCTGCAGAGACGGGACCGCTGATCTACAAGGTG 369  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
370 AAGTTCGCGGACCAACTTCCCCCGACGGCCCCCGTAATGCAGAAAGAACCATGGGC 429  
Db 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162  
430 TGGAGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGGTGAAAGGCGAGATCCAC 489  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
490 CAGGCCCTGAAGCTGAAGGACGGGGCCACTACCTGGTGAGTTCAAGACCATCTACATG 549  
Db 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
550 GCCAAGAAGCCCGTGCAAGCTGCCCGCTACTACTACGTGGACACCAAGCTGGAATCAC 609  
QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222  
610 TCCCAACAAGAGACTACACCATCTGTGAACAAGTACGAGCGCTCCGAGGGCGGCCACAC 669  
Db 223 LeuPheLeu 225  
QY 223 LeuPheLeu 225  
Db 670 CTGTTCTCTG 678

RESULT 12  
US-11-218-880-7  
; Sequence 7, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "dimer2"  
US-11-218-880-7

Alignment Scores:  
Pred. No.: 2.15e-122 Length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservative: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0

DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-7 (1-681)  
QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22  
10 TCCTCCGAGGACGTCTCAAGAAGTTTCATCGCTTCAAAGTGCGCATGAGAGGCTCCGTG 69  
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42  
70 AACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAG 129  
Db 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62  
130 ACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCCCTTGCGCTGGACATCTGTCTC 189  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
190 CCCAGTTCACAGTACGGCTCCAAAGCGGTACGTGAAGCACCCCGGACATCCCGGACTAC 249  
QY 83 LysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGluAspGly 102  
250 AAGAAGCTGTCTCTCCCGGAGGCTTCAAGTGGAGCGCGGTGATGAATCTCGAGACGGC 309  
Db 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
310 GCGGTGGTGACCGTGACCCAGGACTCTCCCTGCAGAGACGGGACCGCTGATCTACAAGGTG 369  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
370 AAGTTCGCGGACCAACTTCCCCCGACGGCCCCCGTAATGCAGAAAGAACCATGGGC 429  
Db 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162  
430 TGGAGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGGTGCTGAAGGGCGAGATCCAC 489  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
490 CAGGCCCTGAAGCTGAAGGACGGGCCCACTACTGTGAGTTCAAGACCATCTACATG 549  
Db 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
550 GCCAAGAAGCCCGTGCAAGCTGCCCGCTACTACTACGTGGACACCAAGCTGGAATCAC 609  
QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222  
610 TCCCAACAAGAGACTACACCATCTGTGAACAAGTACGAGCGCTCCGAGGGCGGCCACAC 669  
Db 223 LeuPheLeu 225  
QY 223 LeuPheLeu 225  
Db 670 CTGTTCTCTG 678

RESULT 13  
US-10-209-208-9  
; Sequence 9, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0



```

; SEQ ID NO 9
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "mRFP1"
US-10-209-208-9

```

```

Alignment Scores:
Pred. No.:      1.36e-110      Length:      678
Score:          1021.00        Matches:     192
Percent Similarity: 89.14%      Conservative: 5
Best Local Similarity: 86.88%      Mismatches: 24
Query Match:     84.10%        Indels:      0
DB:              6            Gaps:           0

```

US-10-006-922A-12 (1-225) x US-10-209-208-9 (1-678)

```

QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 ATGGCCTCTCCGAGGACGTATCATCAAGAGTTCAATCGCTTCAAGGTCCGCATGGAGGCG 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
        ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      61 TCCGTGAACGGCCACGAGTTGAGATCGAGGCGGAGGCGGAGGCGCCCTACGAGGGC 120

QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGCGGCCCTCGCCCTTCGCTGGGACATC 180

```

```

QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      181 CTGTCCCTCAGTTCAGTACGGCTCCAGGCTTACGTGAAGCACCCGCGCATCCCC 240

QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      241 GACTACTTGAAAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGAATTGACTTCGAG 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      301 GACGGCGGCGTGTGACCGTGACCCAGGACTCTCTCCGAGGACCGCGGAGTTCAATCTAC 360

QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      361 AAGGTGAAGCTGCGGCGCACCACTTCCCTCCGACGGCCCCGTAAATGCAGAAAGAACCC 420

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      421 ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCCGAGGACCGCGCCTGAAGGGCGAG 480

QY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      481 ATCAAGATGAGGCTGAAGCTGAAGGACGCGCGGCTCAAGACCGGAGGTCAAGACCACC 540

QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      541 TACATGGCCAAAGAGCCCTGCAAGCTGCGCGGCTTACAAGACCGACATCAAGCTGGAC 600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGlnTyrGluArgThrGluGlyArg 220
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      601 ATCACTCCCAACAGAGACTACACCATCGTGAAAGATAAGAGCGCGCGGAGGCGCGC 660

QY      221 His 221
        |||
DB      661 CAC 663

```

```

RESULT 14
US-11-218-880-9
; Sequence 9, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Taisen, Roger

```

```

; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/11/218, 880
; PRIOR APPLICATION NUMBER: 2005-09-01
; PRIOR FILING DATE: 2002-04-10
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "mRFP1"
US-11-218-880-9

```

```

Alignment Scores:
Pred. No.:      1.36e-110      Length:      678
Score:          1021.00        Matches:     192
Percent Similarity: 89.14%      Conservative: 5
Best Local Similarity: 86.88%      Mismatches: 24
Query Match:     84.10%        Indels:      0
DB:              7            Gaps:           0

```

US-10-006-922A-12 (1-225) x US-11-218-880-9 (1-678)

```

QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1 ATGGCCTCTCCGAGGACGTATCATCAAGAGTTCAATCGCTTCAAGGTCCGCATGGAGGCG 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
        ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      61 TCCGTGAACGGCCACGAGTTGAGATCGAGGCGGAGGCGGAGGCGCCCTACGAGGGC 120

QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGCGGCCCTCGCCCTTCGCTGGGACATC 180

QY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      181 CTGTCCCTCAGTTCAGTACGGCTCCAGGCTTACGTGAAGCACCCGCGCATCCCC 240

QY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      241 GACTACTTGAAAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGAATTGACTTCGAG 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      301 GACGGCGGCGTGTGACCGTGACCCAGGACTCTCTCCGAGGACCGCGGAGTTCAATCTAC 360

QY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      361 AAGGTGAAGCTGCGGCGCACCACTTCCCTCCGACGGCCCCGTAAATGCAGAAAGAACCC 420

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      421 ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCCGAGGACGCGCGCTGAAGGGCGAG 480

QY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      481 ATCAAGATGAGGCTGAAGCTGAAGGACGCGCGGCTCAAGACCGGAGGTCAAGACCACC 540

QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      541 TACATGGCCAAAGAGCCCTGCAAGCTGCGCGGCTTACAAGACCGACATCAAGCTGGAC 600

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:04:28 ; Search time 77 Seconds  
(without alignment)  
1283.899 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVIKEFMRFKVRMEG.....EDYTIQEYERTEGRHMLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 109

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : A\_Geneseq\_21:\*

1: geneseqp19808:\*

2: geneseqp19908:\*

3: geneseqp20008:\*

4: geneseqp20018:\*

5: geneseqp20028:\*

6: geneseqp20038:\*

7: geneseqp20038:\*

8: geneseqp20048:\*

9: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1214	100.0	225	3	AAV99836
2	1214	100.0	225	3	AAB01622
3	1214	100.0	225	4	AAG65509
4	1214	100.0	225	5	ABB08834
5	1214	100.0	225	5	AAE28833
6	1214	100.0	225	5	AAE17540
7	1214	100.0	225	5	AAO18270
8	1214	100.0	225	5	ADY51716
9	1214	100.0	225	6	AAE34962
10	1214	100.0	225	7	ADC24126
11	1214	100.0	225	7	ABW00918
12	1214	100.0	225	7	ADF70403
13	1214	100.0	225	7	ADH34489
14	1214	100.0	225	7	ADL46203
15	1214	100.0	225	7	ADN33979
16	1214	100.0	225	8	ADI36421
17	1214	100.0	225	8	ADM97769
18	1214	100.0	225	8	ADQ28780
19	1214	100.0	225	9	ADX26534
20	1214	100.0	487	5	ABB08821
21	1214	100.0	506	5	ABB08822
22	1214	100.0	547	5	ABB08823
23	1211	99.8	225	5	AAE28920
24	1211	99.8	225	7	ABW00937

25	1211	99.8	225	7	ABW00929	Abw00929	Discosoma
26	1211	99.8	225	7	ABW00938	Abw00938	Discosoma
27	1211	99.8	225	7	ADH34498	Adh34498	Discosoma
28	1211	99.8	225	7	ADH34499	Adh34499	Discosoma
29	1210	99.7	225	5	AAE28919	Aae28919	Discosoma
30	1210	99.7	225	7	ABW00930	Abw00930	Discosoma
31	1210	99.7	226	4	AAG65510	Aag65510	Anthozoan
32	1210	99.7	242	7	ADP24109	Adp24109	Discosoma
33	1210	99.7	545	7	ADL18132	Adl18132	RFP:PS (NI
34	1210	99.7	548	7	ADL18156	Adl18156	RFP:PS (HI
35	1209	99.6	225	5	AAE28922	Aae28922	Discosoma
36	1209	99.6	225	7	ABW00936	Abw00936	Discosoma
37	1209	99.6	225	7	ABW00931	Abw00931	Discosoma
38	1209	99.6	226	5	ABB08835	Abb08835	Yeast opt
39	1208	99.5	225	7	ABW00932	Abw00932	Discosoma
40	1208	99.5	225	7	ABW00935	Abw00935	Discosoma
41	1208	99.5	225	7	ABW00939	Abw00939	Discosoma
42	1208	99.5	225	7	ADH34500	Adh34500	Discosoma
43	1207	99.4	225	5	AAE28921	Aae28921	Discosoma
44	1207	99.4	225	5	AAE17541	Aae17541	Discosoma
45	1207	99.4	225	5	ADY51734	Ady51734	Discosoma
46	1207	99.4	225	7	ABW00941	Abw00941	Discosoma
47	1207	99.4	225	7	ABW00940	Abw00940	Discosoma
48	1207	99.4	225	7	ADL46222	Adl46222	Discosoma
49	1206	99.3	225	7	ABW00933	Abw00933	Discosoma
50	1206	99.3	225	7	ABW00934	Abw00934	Discosoma
51	1205	99.3	225	7	ADH34501	Adh34501	Discosoma
52	1205	99.3	240	6	ABP56678	Abp56678	Mammalian
53	1204	99.2	225	5	AAE28923	Aae28923	Discosoma
54	1204	99.2	225	5	AAE28925	Aae28925	Discosoma
55	1202	99.0	225	5	AAE28924	Aae28924	Discosoma
56	1202	99.0	225	7	ADH34503	Adh34503	Discosoma
57	1202	99.0	225	7	ADH34502	Adh34502	Discosoma
58	1202	99.0	240	6	ABP56685	Abp56685	Discosoma
59	1200	98.8	225	5	ABP70039	Abp70039	Colour Fa
60	1199	98.8	225	7	AAE28926	Aae28926	Discosoma
61	1199	98.8	225	5	AAE28925	Aae28925	Discosoma
62	1199	98.8	240	6	ABP56684	Abp56684	Discosoma
63	1199	98.8	240	6	ABP56683	Abp56683	Discosoma
64	1196	98.5	225	7	ADH34504	Adh34504	Discosoma
65	1196	98.5	236	8	ADO78068	Ado78068	Coralimo
66	1196	98.5	236	8	ADO78070	Ado78070	Coralimo
67	1196	98.5	236	9	ADZ84223	Adz84223	Actinodis
68	1196	98.5	236	9	ADZ84225	Adz84225	Actinodis
69	1195.5	98.5	284	8	ADQ59562	Adq59562	Discosoma
70	1194	98.4	240	6	ABP56681	Abp56681	Discosoma
71	1193	98.3	225	7	ADH34506	Adh34506	Discosoma
72	1193	98.3	240	6	ABP56679	Abp56679	Discosoma
73	1192	98.2	240	6	ABP56680	Abp56680	Discosoma
74	1191	98.1	225	5	AAE28837	Aae28837	Discosoma
75	1191	98.1	225	5	AAE17542	Aae17542	Discosoma
76	1188	97.9	240	6	ABP56682	Abp56682	Discosoma
77	1186	97.7	225	5	AAE28836	Aae28836	Discosoma
78	1186	97.7	225	8	ADH10227	Adh10227	Red fluor
79	1186	97.7	225	8	ADH10228	Adh10228	Protein S
80	1186	97.7	239	8	ADT77528	Adt77528	Immobilis
81	1186	97.7	239	8	ADT77528	Adt77528	Protein/1
82	1183	97.4	240	6	ABP56686	Abp56686	Discosoma
83	1175.5	96.8	226	8	ADL26532	Adl26532	Discosoma
84	1160	95.6	225	7	ADC24129	Adc24129	Discosoma
85	1160	95.6	225	7	ADL46206	Adl46206	Discosoma
86	1153	95.0	225	7	ADL46226	Adl46226	Discosoma
87	1132	93.2	225	8	ADM13537	Adm13537	Superfold
88	1121	92.3	226	7	ADC24131	Adc24131	Discosoma
89	1121	92.3	226	7	ADL46208	Adl46208	Discosoma
90	1119	92.2	225	3	AAV99838	Aay99838	Humanised
91	1119	92.2	225	3	AAE28835	Aae28835	Humanised
92	1111	91.5	225	5	AAV99839	Aay99839	Humanised
93	1111	91.5	225	5	AAE28928	Aae28928	Discosoma
94	1106	91.1	223	5	AAE28929	Aae28929	Discosoma
95	1105	91.0	225	3	AAV99840	Aay99840	Humanised
96	1105	91.0	225	5	AAE28927	Aae28927	Humanised
97	1085.5	89.4	230	3	AAV92894	Aay92894	Discosoma



98	1085.5	89.4	230	3	AAB01625	Aab01625	Discosoma
99	1085.5	89.4	230	4	AAE00378	Aae00378	Discosoma
100	1085.5	89.4	230	5	ABP70038	Abp70038	Colour Fa
101	1085.5	89.4	230	6	ABP96651	Abp96651	Red fluor
102	1058	87.1	205	4	AAE00375	Aae00375	Discosoma
103	1054	86.8	205	4	AAE00384	Aae00384	Discosoma
104	1053	86.7	205	4	AAE00383	Aae00383	Discosoma
105	1021	84.1	225	7	ADC24133	Adc24133	Discosoma
106	1021	84.1	225	7	ADL46210	Adl46210	Discosoma
107	1021	84.1	225	8	ADQ59566	Adq59566	Discosoma
108	1021	84.1	747	9	AEA54910	Aea54910	Fluoresce
109	1012	83.4	225	7	ADL46281	Adl46281	Discosoma

ALIGNMENTS

RESULT 1

AA99836  
ID   AA99836   standard; protein; 225 AA.

AC   AA99836;

DT   12-SEP-2003   (revised)  
DT   19-SEP-2000   (first entry)

DE   Discosoma sp. "red" novel fluorescent protein drFP583.

KW   Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW   fluorescent labeling.

OS   Discosoma sp; "red".

XX   Key                   Location/Qualifiers

FT   Misc-difference 122       /note= "encoded by TC"

FT   Misc-difference 127       /note= "encoded by GTG"

FT   Misc-difference 127       /note= "encoded by GTG"

XX   WO200034326-A1.

XX   15-JUN-2000.

XX   10-DEC-1999;   99WO-US029473.

XX   11-DEC-1998;   98US-00210330.

PR   14-OCT-1999;   99US-00418529.

XX   (CLON-) CLONTECH LAB INC.

PI   Lukyanoy SA,   Fradkov AF,   Labas YA,   Matz MV,   Green G,   Chen Y;  
PI   Ding L;

XX   WPI; 2000-423381/36.

PT   Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT   useful for fluorescent labeling and as markers.

PS   Claim 20; Page 74-75; 86pp; English.

XX   The present sequence is a novel fluorescent protein (nFP) encoded by the  
CC   full-length cDNA drFP583. drFP583 was isolated from Discosoma sp. "red",  
CC   a non-bioluminescent species of the Class Anthozoa. Fluorescent proteins  
CC   can be used in fluorescent labeling, a useful tool for marking a protein,  
CC   cell or organism of interest. Unlike other markers used in protein  
CC   labeling, such as beta-galactosidase and luciferase, fluorescent proteins  
CC   do not require an exogenous cofactor or substrate. Methods involving  
CC   fluorescent proteins are also less laborious and less difficult to  
CC   control than the traditional methods of fluorescent labeling, where a  
CC   protein of interest is purified and then covalently conjugated to a  
CC   fluorophore derivative. Novel fluorescent proteins isolated from species  
CC   of the Class Anthozoa can be used as markers for gene expression and  
CC   protein localization studies, and in fluorescence resonance energy

CC   transfer (FRET) reactions. They may have improved properties and better  
CC   suitability for larger excitations compared to prior art fluorescent  
CC   proteins such as green fluorescent protein. (Updated on 12-SEP-2003 to  
CC   standardise OS field)

XX   SQ   Sequence 225 AA;

Query Match                   100.0%;   Score 1214;   DB 3;   Length 225;  
Best Local Similarity   100.0%;   Pred. No. 1e-127;  
Matches   225;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

QY   1   MRSSKNVLIKFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI   60

DB   1   MRSSKNVLIKFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI   60

QY   61   LSPQFGYSKYVYKHPADIPDYKKLSFPEGFKMERVMNFEDGVTVTQDSSLQDGCFTY   120

DB   61   LSPQFGYSKYVYKHPADIPDYKKLSFPEGFKMERVMNFEDGVTVTQDSSLQDGCFTY   120

QY   121   KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGHYLVFCKSI   180

DB   121   KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGHYLVFCKSI   180

QY   181   YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL   225

DB   181   YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL   225

RESULT 2

AAB01622  
ID   AAB01622   standard; protein; 225 AA.

AC   AAB01622;

DT   12-DEC-2000   (first entry)

DE   Discosoma sp. red fluorescent protein drFP583.

KW   Anthozoa; fluorescent protein; fluorescence labeling; amFP486; cFP484;  
KW   zFP506; zFP538; drFP583; dsFP483; asFP600; dgFP512; dmFP592.

XX   Discosoma sp.

XX   WO200034526-A1.

XX   15-JUN-2000.

XX   10-DEC-1999;   99WO-US029405.

PR   11-DEC-1998;   98US-00210330.

XX   (CLON-) CLONTECH LAB INC.

PI   Lukyanoy SA,   Fradkov AF,   Labas YA,   Matz MV;

XX   WPI; 2000-423451/36.

PT   Novel method for identifying a DNA sequence encoding fluorescent proteins  
PT   from non-bioluminescent Anthozoa which are useful for fluorescent  
PT   labeling and as markers.

PS   Claim 3; Page 68-69; 73pp; English.

XX   The present sequence is Discosoma sp. red fluorescent protein drFP583. It  
CC   was isolated using the Aequoria victoria green fluorescent protein (GFP)  
CC   sequence, which was used to design PCR primers which might isolate other  
CC   fluorescent proteins from a number of species of Anthozoa. These were  
CC   Anemonia majano, Clavularia sp., Zoanthus sp., Discosoma sp. and Anemonia  
CC   sulcata. The cDNA obtained was then screened in the search for sequences  
CC   encoding fluorescent proteins. The other proteins found in this manner  
CC   were cFP484, zFP506, zFP538, amFP486, dsFP483, asFP600, dgFP512 and  
CC   dmFP592. These proteins can be used as fluorescent labels (for gene  
CC   expression and protein localisation studies and in fluorescence resonance



CC energy transfer (FRET) studies) in place of fluorophore derivatives and  
CC luciferases, as these involve laborious processes and the latter require  
CC cofactors. They can also be used in place of GFP, which is too stable to  
CC be useful when studying short-term or repetitive events  
XX

Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 3; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
OY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
OY 121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
Db 121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 3  
AAG65509 standard; protein; 225 AA.

ID AAG65509 standard; protein; 225 AA.  
XX AAG65509;  
AC AAG65509;  
DT 30-NOV-2001 (first entry)

XX Anthozoan red fluorescent protein sequence.

DE Fluorescent protein; Anthozoan; fluorescence; marker; FRET; red.

KW Anthozoa.

OS WO200162919-A1.

XX 30-AUG-2001.

PD 13-FEB-2001; 2001WO-US004625.

XX 23-FEB-2000; 2000US-0184732P.

PR (AURO-) AURORA BIOSCIENCES CORP.

XX Nelson D, Zamaira E, Tsien R;

PI WPI; 2001-557704/62.

XX Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
PT functional red fluorescent proteins, and the encoding nucleic acids, with  
PT key mutations for improving the proteins function.  
XX

PS Disclosure; Page 85; 90pp; English.

XX The invention provides a nucleic acid encoding functional red fluorescent  
CC protein (II) that differs from the sequence of an Anthozoan red  
CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used as a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aequorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,

CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents an anthozoan fluorescent protein  
XX

Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 4; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
OY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
OY 121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
Db 121 KVKEIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGGEIHKALKKDGGHYLVFEKSI 180  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 4  
ABB08834 standard; protein; 225 AA.

ID ABB08834 standard; protein; 225 AA.  
XX ABB08834;  
AC ABB08834;  
DT 29-MAY-2002 (first entry)

XX Yeast optimised RFP SEQ ID NO 17.

DE Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology.  
XX

OS Anthozoa.

XX DE20001395-U1.

PN 15-MAR-2001.

PD 27-JAN-2000; 2000DE-02001395.

XX 27-JAN-2000; 2000DE-02001395.

XX (GPCB-) GPC BIOTECH AG.

XX WPI; 2002-228394/29.

DR N-PSDB; ABA95905, ABA95921, ABA95922.

XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX

PS Disclosure; Page 13-14; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (YRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present



CC	sequence is that of the yeast optimised RFP
XX	
50	Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRSSKAVIKEFMRFKVRMEGTIVNGHEFEIEGEGEGRPYEGHNTVKLKVTYKGBLPFPAWDI	60
Dp	1	MRSSKAVIKEFMRFKVRMEGTIVNGHEFEIEGEGEGRPYEGHNTVKLKVTYKGBLPFPAWDI	60
Qy	61	LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFIY	120
Dp	61	LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFIY	120
Qy	121	KVKFIVGNFPSPDGPVMQKKTMGWEASTERLYPRDVLKGEIHKALKLKDGGHYLVFEKSI	180
Dp	121	KVKFIVGNFPSPDGPVMQKKTMGWEASTERLYPRDVLKGEIHKALKLKDGGHYLVFEKSI	180
Qy	181	YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL	225
Dp	181	YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL	225

## RESULT 5

AAAE28833

**ID** AAE28833 standard; protein; 225 AA.

AC AAE28833;

DT 27-DEC-2002 (first entry)

DE Discosoma sp. drFP583 (NFP-6) wild-type protein.

KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW fluorescence activated cell sorting application; fluorescent timer;  
 KW biosensor; fluorescence resonance energy transfer application; FRET;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW bunscreen; second messenger detector; drpP583 protein; NFP-6.

**OS Diacosoma sp.**

PN W0200268459-A2.

PD 06-SEP-2002.

PF 20-FEB-2002; 2002WO-US005749.

PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.

PA (CLON-) CLONTECH LAB INC.

PI Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;

DR WPI; 2002-691654/74.  
DR N-PSDB; AAD46278.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications

PS Disclosure; Page 70-71; 80pp; English.

The invention relates to nucleic acid molecules encoding non-aggregating chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are useful in analyte detection assays, as colouring agents, as markers in recombinant DNA applications, as sunscreens or filters, in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in screening assays, as second messenger detectors, in fluorescence activated cell sorting applications,

CC in protease cleavage assays or as fluorescent trimers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) wild-type protein of the  
CC invention

**SQ** sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MSSSKNVIKEFMRFKVMMEGTIVNGHFEIIEGEGEGRPYEGHNTVTKLVTKKGSPLEPFAMDI	60
Dp	1	MSSSKNVIKEFMRFKVMMEGTIVNGHEFEIEGEGEGRPYEGHNTVKLKVTKKGSPLEPFAMDI	60
QY	61	LSPQFOYGSKVYKHHPADIPDYKLSFPEGFKWERVMNFEDGVTVTQDSSLDDCGFIY	120
Dp	61	LSPQFOYGSKVYKHHPADIPDYKLSFPEGFKWERVMNFFEDGVTVTQDSSLDDCGFIY	120
QY	121	KVFFIGVNFPSDGPVMOCKTMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI	180
Dp	121	KVFFIGVNFPSDGPVMOCKTMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEEKSI	180
QY	181	YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTBGRHHLFL	225
Dp	181	YMAKKPVQLPGYYVYDSKLDITSHNEDYTIIVEQYERTBGRHHLFL	225

## RESULT 6

AAAE17540

**ID** AAE17540 standard; protein; 225 AA.

AC      AAE17540;

DT 22-APR-2002 (First entry)

DE Discosoma sp. humanised wild-type Anthozoa protein drFP583.

KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; anthozoa protein; drfp583.

**OS Discosoma sp.**

PN WO200196373-A2.

PD 20-DEC-2001.

PF 13-JUN-2001; 2001WO-US019097.

PR 14-JUN-2000; 2000US-0211607P.

PA (CLON-) CLONTECH LAB INC.

PI Fradkov AF, Tersikh A;

DR WPI; 2002-154595/20.

PT New fluorescent timer proteins comprising an emission spectrum that changes over time from a first wavelength to a second wavelength, useful for monitoring intracellular protein movement, translocation, trafficking or stability.

PS Example 1; Fig 1; 89pp; English.

The invention relates to a fluorescent timer protein having an emission spectrum that changes over time after synthesis from a first wavelength to a second wavelength. The fluorescent timer proteins are useful in monitoring the activity of a promoter, determining the age of a protein identifying an agent that modulates the activity of a promoter and in enriching a population of cells comprising a fluorescent timer protein. The fluorescent timer proteins are also useful for assessing gene expression during development of a multicellular organism or during



cellular differentiation, in response to a drug or other inducer of promoter activity, as a reporter to serve as a read-out of promoter activity, monitoring intracellular protein movement or translocation, protein trafficking, or protein stability, to investigate temporal aspects of the activity of a regulatory element, for determining cell fate during development and organ remodelling, in spatial and temporal visualisation of newly synthesised proteins and accumulated proteins, and in distinguishing between newly formed and pre-existing structures, e.g. membrane junctions and extracellular matrix components. The fluorescent timer proteins may further be used to investigations where photobleaching techniques are employed, as detectable labels, as selectable markers, as biosensors in prokaryotic and eukaryotic cells, in protease cleavage assays, and as second messenger detectors. The nucleic acids can be used to generate transgenic, non-human plants or animals or site-specific gene modifications in cell lines. The present sequence is *Drosophila* sp. humanised wild-type Anthozoa protein dRFP583 used for generating fluorescent proteins

Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

## RESULT 7

ID AA018270 standard; protein; 225 AA.

AC AA018270;

DT 26-SEP-2002 (first entry)

Discosoma red fluorescent protein.

Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette; modified yeast strain; environmental pollution.

**OS Discovoma sp.**

PN DE10061872-A1.

PD 20-JUN-2002.

12-DEC-2000; 2000DE-01061872.

PR 12-DEC-2000, 2000DE-01061872.

PA (LICH/) LICHTENBERG-PRATE H.

PI Lichtenberg-Prate H;

DR WPI; 2002-539633/58.

DR N-PSDB; AAL47952.

Modified yeast strain, useful for detecting toxic compounds in PT environment, contains integrated cassettes responsive to genotoxic and

PT cytotoxic compounds.

PS Disclosure; Page 21-22; 34pp; German.

The present invention relates to a modified yeast strain that contains, integrated stably and functionally in its genome, a genotoxicity cassette and a cytotoxicity cassette, each comprising a promoter and reporter gene, both of which are different in the two cassettes. The modified yeast strain is used to detect environmental pollution, especially genotoxic and/or cytotoxic substances in complex environmental contaminants, especially organic compounds, but also (non-)ionising radiation and chemical carcinogens. Particular applications are in monitoring (waste) water (e.g. as an early warning system), medical toxicology screening and for industrial process control. The present sequence is a marker protein suitable for use in the cassettes of the present invention

Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225; Conservative	0;	Mismatches	0;	Gaps 0;

QY		1	MRSSKNVKEFMRFKVRMEGTIVNGHFEIEGEGERPRYEGHNTVKLKVTKGGLPLPFAMDI	60
Dd		1	MSSSKNVIKEFMRFKVRMEGTIVNGHBEIEGEGGRPYEGHNTVLKVTKGGPLPFAWDI	60
QY		61	LSPFOYGSKVVYVKHPADIPDYKKLSFPEGFWERVMNFEDGVTVTQDSSLQDGCFIY	120
Dd		61	LSPFOYGSKVVYVKHPADIPDYKKLSFPEGFKWERMVNFFEDGVTVTITQDSSLQDGCFIY	120
QY		121	KVFFIGVNPSPDGPVMÖKTMTGWASTERTLYPRDGLKGEIHKALKLKDGSHTLVVEFKSI	180
Dd		121	KVFFIGVNPSPDGPVMÖKTMTGWASTERTLYPRDGLKGEIKHALKLKDGSHTLVVEFKSI	180
QY		181	YMAKKPVQLPGYYVDSDLDTSHNEDYTIVEOYERTEGRHHLLFL	225
Dd		181	YMAKKPVQLPGYYVDSDKLDITSHNEDYTIVEOYERTEGRHHLLFL	225

## RESULT 8

ID ADY51716 standard; protein; 225 AA.

AC ADY51716;

DT 05-MAY-2005 (first entry)

DE *Discosoma* sp. wild type red fluorescent protein (RFP) Seq 12.

fluorescence; mutagenesis; red fluorescent protein; protein interaction.

OS *Discosoma* sp.

PN W0200268605-A2.

06-SEP-2002  
PD

26-FEB-2002; 2002WO-US006063.

26-FEB-2001; 2001US-00794308.

XX






















XX

DR N-PSDB; ADY51715.

XX

aa New non-oligomerizing fluorescent protein containing at least one  
PT mutation that reduces or eliminates the ability of the protein to  
PT oligomerize, useful for making better and new assays for molecular







PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.

PS Claim 1; SEQ ID NO 1; 67bp; English.

CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (i) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localization or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma wild-type red fluorescent protein.

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
OY 61 LSPQFGYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 11

ABW00918 ID ABW00918 standard; protein; 225 AA.

XX AC ABW00918;

XX DT 15-JAN-2004 (first entry)

XX DE Discosoma sp. red fluorescent protein (RFP).

XX KM Fluorescent protein; resonance energy transfer; pH; detection;

KW red fluorescent protein; RFP.

XX OS Discosoma sp.

XX PN US2003170911-A1.

XX PD 11-SEP-2003.

XX PF 26-FEB-2001; 2001US-00794308.

XX PR 26-FEB-2001; 2001US-00794308.

XX PA (TSIE/) TSIE R Y.

PA (ZACH/) ZACHARIAS D A.

PA (BAIR/) BAIRD G S.

XX PI Tsien RY, Zacharias DA, Baird GS;

XX DR WPI; 2003-802418/75.

DR N-PSDB; AAD61969.

PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.

PS Claim 10; Page 30-31; 0pp; English.

CC The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerize. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP)

XX SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
OY 61 LSPQFGYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 12

ADF70403 ID ADF70403 standard; protein; 225 AA.

XX AC ADF70403;

XX DT 12-FEB-2004 (first entry)

XX DE Discosoma wild-type GFP variant protein SeqID26.

KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;

KW GFPuv; Enhanced GFP; EGFP.

XX OS Discosoma sp.

XX PN WO2003071272-A1.







CC represents a wild-type red fluorescent protein, DsRed, from Discosoma sp.  
CC that was used as a parent sequence for the generation of mutant proteins  
CC in an example of the invention.

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAMD 60  
1 MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAMD 60  
DB 1 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 14

ADL46203  
ID ADL46203 standard; protein; 225 AA.

XX AC ADL46203;  
XX DT 20-MAY-2004 (first entry)  
XX DE Discosoma red fluorescent (DsRed) protein.

XX DE Discosoma red fluorescent (DsRed) protein; DsRed; fluorescence; red wavelength;  
XX KW oligomerization; tetramerization; immunoassay; hybridization assay.

XX OS Discosoma sp.  
XX PN WO2003086446-A1.  
XX PD 23-OCT-2003.  
XX PF 09-APR-2003; 2003WO-US010879.  
XX PR 10-APR-2002; 2002US-00121258.  
XX PR 29-JUL-2002; 2002US-00209208.

XX PA (REGC ) UNIV CALIFORNIA.  
XX PI Telen RY, Campbell RE, Baird GS;  
XX DR WPI; 2003-845265/78.  
XX DR N-PSDB; ADL46204.

XX PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
XX PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
XX PT in molecular biology, e.g. in immunoassays or in tracking protein  
XX PT movement in cells.

XX PS Claim 1; SEQ ID NO 1; 166pp; English.

XX CC The invention relates to a polynucleotide sequence encoding a Discosoma  
XX CC red fluorescent protein (DsRed) variant having a reduced propensity to  
XX CC oligomerize. The protein variant comprises one or more amino acid  
XX CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
XX CC sequence, where the substitutions result in reduced propensity of the  
XX CC DsRed variant to form tetramers and where the variant displays detectable  
XX CC fluorescence of at least one red wavelength. The composition and methods  
XX CC are useful in producing red fluorescent proteins having reduced

CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC protein.

XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAMD 60  
1 MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAMD 60  
DB 1 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGYLKGEIHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 15

ADN33979  
ID ADN33979 standard; protein; 225 AA.

XX AC ADN33979;  
XX DT 01-JUL-2004 (first entry)  
XX DE Wild-type DsRED protein.

XX DE Wild-type DsRED protein.  
XX KW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRED.

XX OS Discosoma sp.  
XX PN WO2003054158-A2.  
XX PD 03-JUL-2003.  
XX PF 18-DEC-2002; 2002WO-US040539.  
XX PR 19-DEC-2001; 2001US-0341723P.

XX PA (UYCH-) UNIV CHICAGO.  
XX PI Bevis B, Glick B;  
XX DR WPI; 2003-569236/53.  
XX DR N-PSDB; ADN33978.

XX PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
XX PT mutant of a Cnidarian chromo- or fluorescent protein or its mutant,  
XX PT useful for applications involving chromo- or fluorescent proteins.

XX PS Claim 8; SEQ ID NO 2; 65pp; English.

XX CC The present invention relates to nucleic acid that encodes a rapidly  
XX CC maturing chromo- or fluorescent mutant of a Cnidarian chromo- or  
XX CC fluorescent protein or its mutant. The protein is useful in applications  
XX CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
XX CC useful for producing a chromo- or fluorescent protein which involves  
XX CC growing the cell, whereby the protein is expressed, and isolating the  
XX CC protein substantially free of other proteins. The protein is useful in  
XX CC applications involving chromo- or fluorescent protein and is useful as  
XX CC PCR primers, hybridization probes, etc. The expression cassettes are



CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for  
CC biological analytes of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins  
CC also find use in protease cleavage assays. The proteins can also be used  
CC is assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRED.

XX Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 7; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTGKGPPLPFAWDI 60  
|||  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTGKGPPLPFAWDI 60  
Qy 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
|||  
Db 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
Qy 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYYVVD SKLDTSHNEDYTIVEQYERTEGRHHLFL 225  
|||  
Db 181 YMAKKPVQLPGYYVVD SKLDTSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 16

ID ADI36421  
ID ADI36421 standard; protein; 225 AA.

AC ADI36421;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RED).  
XX  
KW Fluorobody; binding ligand; green fluorescent protein; GFP;  
KW target detection; red fluorescent protein; RED.  
XX  
OS Discosoma sp.

XX Key Location/Qualifiers  
FH 22. .26  
FT Region /note= "Loop region"  
FT 37. .39  
FT Region /note= "Loop region"  
FT 75. .81  
FT Region /note= "Loop region"  
FT 100. .103  
FT Region /note= "Loop region"  
FT 114. .117  
FT Region /note= "Loop region"  
FT 153. .156  
FT Region /note= "Loop region"  
FT 167. .170  
FT Region /note= "Loop region"  
FT 185. .192

FT /note= "Loop region"  
FT Region 204. .209  
FT /note= "Loop region"

PN US2003203355-A1.  
XX  
PD 30-OCT-2003.  
XX  
PF 24-APR-2002; 2002US-00132067.  
XX  
PR 24-APR-2002; 2002US-00132067.  
XX  
PA (LALA-) LOS ALAMOS NAT LAB.  
XX (REGC ) UNIV CALIFORNIA.  
PI Bradbury AM, Zeytun A, Waldo GS;  
XX  
DR WPI; 2004-154325/15.  
DR N-PSDB; ADI36420.  
XX  
PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.  
XX  
PS Claim 2; SEQ ID NO 4; 23pp; English.  
XX  
CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) used in  
CC the exemplification of the invention.  
XX  
SQ Sequence 225 AA;

Query Match 100.0%; Score 1214; DB 8; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTGKGPPLPFAWDI 60  
|||  
Db 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTGKGPPLPFAWDI 60  
Qy 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
|||  
Db 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
Qy 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYYVVD SKLDTSHNEDYTIVEQYERTEGRHHLFL 225  
|||  
Db 181 YMAKKPVQLPGYYVVD SKLDTSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 17

ID ADM97769  
ID ADM97769 standard; protein; 225 AA.

AC ADM97769;  
XX  
DT 01-JUL-2004 (first entry)  
XX  
DE D sp red fluorescent protein SEQ ID NO: 22.  
XX  
KW enzyme; sensor cell; green fluorescent protein;  
KW signal transduction detection system; promoter; targeting sequence;  
KW targeted drug.

XX Discosoma sp.  
OS  
XX WO2004031415-A2.  
PN



XX 15-APR-2004.  
PD  
XX 05-SEP-2003; 2003WO-US028078.  
PF  
XX 05-SEP-2002; 2002US-0408297P.  
PR  
XX  
PA (VERT-) VERTEX PHARM INC.  
XX  
PI Whitney MA, Zeh K, Sanders PS;  
XX  
XX WPI; 2004-330208/30.  
DR N-PSDB; ADM97768.  
DR  
XX  
PT Developing a sensor cell, useful in determining the activity of a target  
PT gene and in developing therapeutic drugs, comprises providing cells  
PT comprising a signal transduction detection system and introducing DNA  
PT construct into cells.  
XX  
PS Disclosure; Page 168-169; 234pp; English.  
XX  
XX The present invention relates to a method of developing a sensor cell,  
CC for determining the activity of a target gene in the cell, which  
CC comprises providing a homogeneous population of cells, where each of the  
CC cells comprises a signal transduction detection system and introducing  
CC into the population of cells an isolated DNA construct comprising a  
CC promoter operatively linked to a targeting sequence. The method is useful  
CC in developing a sensor cell for determining the activity of a target gene  
CC in the cell. The sensor cell and the methods are useful in developing new  
CC and therapeutic drugs directed to the targets. The present sequence is a  
XX polypeptide shown in the exemplification of the invention.  
SQ Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 8;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

RESULT 18  
ADQ28780  
ID ADQ28780 standard; protein; 225 AA.

ADQ28780;	
AC	
XX	
DT	07-OCT-2004 (first entry)

DE Jellyfish green fluorescent protein DsRed.

Green fluorescent protein; fluorescent protein.

*Aegoreea victoria*.

PN WO2004058973-A1.

15-JUL-2004.

XX

PF 26-NOV-2003; 2003WO-RU0005525.  
 XX 26-DEC-2002; 2002US-0436857P.  
 PR 02-APR-2003; 2003US-0459679P.  
 XX (EVRO=) EVROGEN STOCK CO.  
 PA  
 XX Barsova EV, Lukyanov SA;  
 PI  
 XX WPI; 2004-525887/50.  
 DR  
 XX  
 PT New isolated nucleic acid molecule encoding a fluorescent protein, useful  
 PT for labeling biomolecules, cell, or cell organelles, or for identifying  
 PT expression of a gene in a biological specimen.  
 XX  
 XX  
 PS Example 1; Fig 1; 55pp; English.  
 XX  
 CC The present invention relates to novel coding sequences (I) encoding  
 CC green fluorescent proteins (II). The fluorescent protein coding sequences  
 CC are isolated from an organism from phylum Arthropoda, from subclass  
 CC Copepoda, or from family Pontellidae. The fluorescent protein coding  
 CC sequences are useful for labelling biomolecules, cell, or cell  
 CC organelles. They are also useful for identifying expression of a gene in  
 CC a biological specimen or for generating transformants including  
 CC transgenic organisms or site-specific gene modifications in cell lines.  
 CC The fluorescent proteins are also useful in high throughput screening  
 CC assays. The present sequence was used in a sequence alignment with the  
 CC green fluorescent proteins of the invention.  
 XX  
 XX  
 SQ Sequence 225 AA;

Query Match	100.0%;	Score 1214;	DB 8;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 1e-127;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	MRSSKNV	KEFMR	FKVRMEGT	VNGHEFEI	EGEGEG	RPYEGH	NTVKL	KVTKG	BPLP	PAWDI	60
Db	1	MRSSKNV <td>KEFMR <td>FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td></td></td>	KEFMR <td>FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td></td>	FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td>	VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td>	EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td>	RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td>	NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td>	KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td>	BPLP <td>PAWDI</td> <td>60</td>	PAWDI	60
Qy	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK</td> <td>KLSP</td> <td>PEGFK</td> <td>WERVMN</td> <td>FEDG</td> <td>VVTV</td> <td>ODSSL</td> <td>ODGCFIY 120</td> </td>	GSKVYV <td>KHPADI</td> <td>PDYK</td> <td>KLSP</td> <td>PEGFK</td> <td>WERVMN</td> <td>FEDG</td> <td>VVTV</td> <td>ODSSL</td> <td>ODGCFIY 120</td>	KHPADI	PDYK	KLSP	PEGFK	WERVMN	FEDG	VVTV	ODSSL	ODGCFIY 120
Db	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK</td> <td>KLSP</td> <td>PEGFK</td> <td>WERVMN</td> <td>FEDG</td> <td>VVTV</td> <td>ODSSL</td> <td>ODGCFIY 120</td> </td>	GSKVYV <td>KHPADI</td> <td>PDYK</td> <td>KLSP</td> <td>PEGFK</td> <td>WERVMN</td> <td>FEDG</td> <td>VVTV</td> <td>ODSSL</td> <td>ODGCFIY 120</td>	KHPADI	PDYK	KLSP	PEGFK	WERVMN	FEDG	VVTV	ODSSL	ODGCFIY 120
Qy	121	KVKF	IGVNF	PSDGP	VMO	KTMG	WEAS	TERL	YPRD	GV	LKGEI	HKAL
Db	121	KVKF	IGVNF	PSDGP	VMO	KTMG	WEAS	TERL	YPRD	GV	LKGEI	HKAL
Qy	181	YMAK	KPVOL	PGY	YVDS	KLDIT	SHNE	DTIVE	QYER	T	EGRH	LFL 225
Db	181	YMAK	KPVOL	PGY	YVDS	KLDIT	SHNE	DTIVE	QYER	T	EGRH	LFL 225

RESULT 19  
ADPX26534  
ID ADPX26534 standard; protein; 225 AA.

AC ADX26534;

MM 21-APR-2005 (first entry)  
DT

Discosoma DAREd RFP protein, seq id 12.

phosphorylation; detection; red fluorescent protein; RFP.

**Discosoma sp.**

US2005026234-A1.

PD 03-FEB-2005.

28-MAY-2004; 2004US-00857622.

PR 31-JAN-1996; 96US-00594575.

...

CC absorbance and fluorescence of the invention.  
CC green fluorescent proteins of the invention.  
XX  
SQ Sequence 225 AA;

Qy	1	MRSSKNV	KEFMR	FKVRMEGT	VNGHEFEI	EGEGEG	RPYEGH	NTVKL	KVTKG	BPLP	PAWDI	60
Db	1	MRSSKNV <td>KEFMR <td>FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td></td></td>	KEFMR <td>FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td></td>	FKVRMEGT <td>VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td></td>	VNGHEFEI <td>EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td></td>	EGEGEG <td>RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td></td>	RPYEGH <td>NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td></td>	NTVKL <td>KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td></td>	KVTKG <td>BPLP <td>PAWDI</td> <td>60</td> </td>	BPLP <td>PAWDI</td> <td>60</td>	PAWDI	60
Qy	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td></td>	GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td>	KHPADI	PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td>	KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td>	PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td>	KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td>	VMN	FEDG <td>VTVT</td> <td>ODSSLQDGC</td>	VTVT	ODSSLQDGC
Db	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td></td>	GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td>	KHPADI	PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td>	KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td>	PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td>	KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td>	VMN	FEDG <td>VTVT</td> <td>ODSSLQDGC</td>	VTVT	ODSSLQDGC
Qy	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td></td>	GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td>	KHPADI	PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td>	KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td>	PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td>	KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td>	VMN	FEDG <td>VTVT</td> <td>ODSSLQDGC</td>	VTVT	ODSSLQDGC
Db	61	LSPOFOY <td>GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td></td>	GSKVYV <td>KHPADI</td> <td>PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td></td>	KHPADI	PDYK <td>KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td></td>	KLSP <td>PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td></td>	PEGF <td>KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td></td>	KMER <td>VMN</td> <td>FEDG <td>VTVT</td> <td>ODSSLQDGC</td> </td>	VMN	FEDG <td>VTVT</td> <td>ODSSLQDGC</td>	VTVT	ODSSLQDGC
Qy	121	KVKF <td>IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td></td>	IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td>	PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td>	VMO	KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td>	WEAS	TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td>	YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td>	GV	LKGEI <td>HKAL</td>	HKAL
Db	121	KVKF <td>IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td></td>	IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td>	PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td>	VMO	KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td>	WEAS	TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td>	YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td>	GV	LKGEI <td>HKAL</td>	HKAL
Qy	121	KVKF <td>IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td></td>	IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td>	PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td>	VMO	KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td>	WEAS	TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td>	YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td>	GV	LKGEI <td>HKAL</td>	HKAL
Db	121	KVKF <td>IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td></td>	IGVNF <td>PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td></td>	PSDGP <td>VMO</td> <td>KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td></td>	VMO	KTMG <td>WEAS</td> <td>TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td></td>	WEAS	TERL <td>YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td></td>	YPRD <td>GV</td> <td>LKGEI <td>HKAL</td> </td>	GV	LKGEI <td>HKAL</td>	HKAL
Qy	181	YMAK <td>KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td></td>	KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td>	PGY	YVDS	KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td>	SHNE	DTIVE	QYER	T	EGRH	LFL
Db	181	YMAK <td>KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td></td>	KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td>	PGY	YVDS	KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td>	SHNE	DTIVE	QYER	T	EGRH	LFL
Qy	181	YMAK <td>KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td></td>	KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td>	PGY	YVDS	KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td>	SHNE	DTIVE	QYER	T	EGRH	LFL
Db	181	YMAK <td>KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td></td>	KPVOL <td>PGY</td> <td>YVDS</td> <td>KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td> </td>	PGY	YVDS	KLDIT <td>SHNE</td> <td>DTIVE</td> <td>QYER</td> <td>T</td> <td>EGRH</td> <td>LFL</td>	SHNE	DTIVE	QYER	T	EGRH	LFL

RESULT 19  
ADPX26534  
ID ADPX26534 standard; protein; 225 AA.

AC ADX26534;

MM 21-APR-2005 (first entry)  
DT

Discosoma DAREd RFP protein, seq id 12.

phosphorylation; detection; red fluorescent protein; RFP.

**Discosoma sp.**

US2005026234-A1.

PD 03-FEB-2005.

28-MAY-2004; 2004US-00857622.

PR 31-JAN-1996; 96US-00594575.

...







XX WO200212543-A2.  
PN  
XX  
PD 14-FEB-2002.  
XX  
PF 07-AUG-2001, 2001WO-EP009112.  
XX  
PR 07-AUG-2000, 2000DE-01038382.  
XX  
PA (DIRE-) DIREVO BIOTECH AG.  
XX  
PI Kuhlmann R, Koltermann A, Ketting U, Schwille P;  
XX WPI; 2002-269094/31.  
DR  
XX  
XX New autofluorescent fusion protein, useful for determining protease and  
PT protease-inhibiting activity, comprises two different proteins linked by  
PT protease cleavage site.  
XX  
PS Claim 6; Page 11; 35pp; German.  
XX  
XX The invention relates to an autofluorescing fusion protein (I, ABB08821-  
CC ABB08823) comprising: (i) a first autofluorescing protein (Ia); (ii) a  
CC segment containing a protease cleavage site; and (iii) at least one  
CC different autofluorescing protein (Ib). Essentially no fluorescent energy  
CC transfer occurs between (Ia) and (Ib). (I) is used for detecting and  
CC quantifying protease (or protease-inhibitory) activity in liquid samples  
CC or cells, particularly for screening-based optimisation (or generation)  
CC of biomolecules with proteolytic activity. (I) can be prepared in  
CC cellular or cell-free systems and makes possible intracellular analysis  
CC of protease activity. Preparation of (I) does not require regioselective  
CC coupling of fluorophores to polypeptides and any selected protease  
CC cleavage site can be incorporated  
XX  
XX Sequence 506 AA;  
SQ

Query Match 100.0%; Score 1214; DB 5; Length 506;  
Best Local Similarity 100.0%; Pred. No. 3.4e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 282 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 341  
QY 61 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWERVMNPFEDGGVVTQTODSSLQDGCFTY 120  
DB 342 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWERVMNPFEDGGVVTQTODSSLQDGCFTY 401  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEFSI 180  
DB 402 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEFSI 461  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 462 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 506

RESULT 22  
ABB08823  
ID ABB08823 standard; protein, 547 AA.  
XX  
AC ABB08823;  
XX  
DT 28-MAY-2002 (first entry)  
XX  
DE Autofluorescent fusion protein C SEQ ID NO 3.  
XX  
KW Autofluorescent; fusion protein; proteolytic activity; fluorophore;  
KW protease; assay; protease cleavage.  
XX  
OS Synthetic.  
XX  
PN WO200212543-A2.

XX 14-FEB-2002.  
PD  
XX  
XX 07-AUG-2001, 2001WO-EP009112.  
PF  
XX  
XX 07-AUG-2000, 2000DE-01038382.  
PR  
XX  
PA (DIRE-) DIREVO BIOTECH AG.  
XX  
PI Kuhlmann R, Koltermann A, Ketting U, Schwille P;  
XX WPI; 2002-269094/31.  
DR  
XX  
XX New autofluorescent fusion protein, useful for determining protease and  
PT protease-inhibiting activity, comprises two different proteins linked by  
PT protease cleavage site.  
XX  
XX Claim 6; Page 11; 35pp; German.  
XX  
XX The invention relates to an autofluorescing fusion protein (I, ABB08821-  
CC ABB08823) comprising: (i) a first autofluorescing protein (Ia); (ii) a  
CC segment containing a protease cleavage site; and (iii) at least one  
CC different autofluorescing protein (Ib). Essentially no fluorescent energy  
CC transfer occurs between (Ia) and (Ib). (I) is used for detecting and  
CC quantifying protease (or protease-inhibitory) activity in liquid samples  
CC or cells, particularly for screening-based optimisation (or generation)  
CC of biomolecules with proteolytic activity. (I) can be prepared in  
CC cellular or cell-free systems and makes possible intracellular analysis  
CC of protease activity. Preparation of (I) does not require regioselective  
CC coupling of fluorophores to polypeptides and any selected protease  
CC cleavage site can be incorporated  
XX  
XX Sequence 547 AA;  
SQ

Query Match 100.0%; Score 1214; DB 5; Length 547;  
Best Local Similarity 100.0%; Pred. No. 3.8e-127;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 323 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 382  
QY 61 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWERVMNPFEDGGVVTQTODSSLQDGCFTY 120  
DB 383 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWERVMNPFEDGGVVTQTODSSLQDGCFTY 442  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEFSI 180  
DB 443 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEFSI 502  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 503 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 547

RESULT 23  
AAE28920  
ID AAE28920 standard; protein, 225 AA.  
XX  
AC AAE28920;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Discosoma sp. drFP583 (NFP-6) mutant protein, E5 (S197T).  
XX  
KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KW mutein.  
XX  
OS Discosoma sp.  
XX



OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 197  
FT /note= "Wild-type Ser substituted with Thr"  
XX  
XX WO200268459-A2.  
XX  
XX 06-SEP-2002.  
XX  
XX 20-FEB-2002, 2002WO-US005749.  
XX  
XX 21-FEB-2001, 2001US-0270983P.  
PR 04-DEC-2001, 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
PI  
XX WPI; 2002-691654/74.  
XX  
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
XX Disclosure; Page; 80pp; English.  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as screens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX Sequence 225 AA;  
SQ  
  
Query Match 99.8%; Score 1211; DB 5; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
  
QY 61 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGGVVTVTQDSSLQDGCFTY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGELHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGELHKALKLKDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
  
RESULT 24  
ABW00937  
ID ABW00937 standard; protein; 225 AA.  
XX  
XX ABW00937;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein (RFP), S197T.  
DE

XX  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; mutein.  
XX  
XX Discosoma sp.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 197  
FT /note= "Wild-type Ser is substituted with Thr"  
XX  
XX US2003170911-A1.  
XX  
XX 11-SEP-2003.  
XX  
XX 26-FEB-2001, 2001US-00794308.  
PF 26-FEB-2001, 2001US-00794308.  
XX  
XX 26-FEB-2001, 2001US-00794308.  
PR  
XX  
XX (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
XX Tsien RY, Zacharias DA, Baird GS;  
PI  
XX WPI; 2003-802418/75.  
XX  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
XX  
XX Example 2; Page; 0pp; English.  
XX  
XX The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX Sequence 225 AA;  
SQ  
  
Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
  
QY 61 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEEDGGVVTVTQDSSLQDGCFTY 120  
  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGELHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLVKGELHKALKLKDGGHYLVEFKSI 180  
  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
  
RESULT 25  
ABW00929  
ID ABW00929 standard; protein; 225 AA.  
XX  
XX ABW00929;  
AC



XX 15-JAN-2004 (first entry)  
XX Discosoma sp. red fluorescent protein (RFP), K83R.  
DE Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; mutant; mutein.  
XX Discosoma sp.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 83 /note= "wild-type Lys is substituted with Arg"  
FT  
FT  
XX US2003170911-A1.  
XX 11-SEP-2003.  
XX 26-FEB-2001; 2001US-00794308.  
XX 26-FEB-2001; 2001US-00794308.  
XX 26-FEB-2001; 2001US-00794308.  
XX (TSIE/) TSIE R Y.  
XX (ZACH/) ZACHARIAS D A.  
XX (BAIR/) BAIRD G S.  
XX Tsien RY, Zacharias DA, Baird GS;  
XX WPI; 2003-802418/75.  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX Example 2; Page; Opp; English.  
XX The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX Sequence 225 AA;  
SQ  
Query Match 99.8%; Score 1211; DB 7; length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

ABW00938  
ID ABW00938 standard; protein; 225 AA.  
XX  
XX AC ABW00938;  
XX 15-JAN-2004 (first entry)  
DT  
XX Discosoma sp. red fluorescent protein (RFP), K70R.  
DE  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; mutant; mutein.  
XX Discosoma sp.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Misc-difference 70 /note= "wild-type Lys is substituted with Arg"  
FT  
FT  
XX US2003170911-A1.  
XX 11-SEP-2003.  
XX 26-FEB-2001; 2001US-00794308.  
XX 26-FEB-2001; 2001US-00794308.  
XX 26-FEB-2001; 2001US-00794308.  
XX (TSIE/) TSIE R Y.  
XX (ZACH/) ZACHARIAS D A.  
XX (BAIR/) BAIRD G S.  
XX Tsien RY, Zacharias DA, Baird GS;  
XX WPI; 2003-802418/75.  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX Example 2; Page; Opp; English.  
XX The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX Sequence 225 AA;  
SQ  
Query Match 99.8%; Score 1211; DB 7; length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 27  
ADH34498  
ID ADH34498 standard; protein; 225 AA.  
XX  
AC ADH34498;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Discosoma sp. DsRed mutant S148A.  
XX  
KW Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labeling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed; mutant; mutein.  
XX  
OS Synthetic.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 148"  
FT  
XX  
XX WO2003057833-A2.  
PN  
PD 17-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US041418.  
PF  
XX  
PR 26-DEC-2001; 2001US-0343128P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
PA  
XX  
PI Bulina ME, Chudakov D, Lukyanov KA;  
XX  
XX WPI; 2003-607998/57.  
DR  
XX  
XX  
PT Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
XX  
PS Example 1; Page; 56pp; English.  
XX  
XX The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs,  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as colouring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and

CC are also useful as sunscreens and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein DsRed mutant generated  
CC in an example of the invention. The present sequence is not shown in the  
CC specification, but was derived from the wild-type DsRed sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.  
XX  
SQ Sequence 225 AA;

Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
|||  
DB 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFDGCVTVTTQDSSLQDGCFTY 120  
|||  
DB 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFDGCVTVTTQDSSLQDGCFTY 120  
QY 121 KVKFTGVNFPSSDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
DB 121 KVKFTGVNFPSSDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 28  
ADH34499  
ID ADH34499 standard; protein; 225 AA.  
XX  
AC ADH34499;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Discosoma sp. DsRed mutant S203A.  
XX  
XX  
KW Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labeling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed; mutant; mutein.  
XX  
OS Synthetic.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 197 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 203"  
FT  
XX  
XX WO2003057833-A2.  
PN  
PD 17-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US041418.  
PF  
XX



PR 26-DEC-2001; 2001US-0343128P.  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Bulina ME, Chudakov D, Lukyanov KA;  
XX  
XX WPI; 2003-607998/57.  
DR  
XX  
XX Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
PS Example 1; Page; 56pp; English.

CC The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs,  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as coloring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and  
CC are also useful as screens and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganization and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein DsRed mutant generated  
CC in an example of the invention. The present sequence is not shown in the  
CC specification, but was derived from the wild-type DsRed sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.

XX Sequence 225 AA;

Query Match 99.8%; Score 1211; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
OY 61 LSPQFQYGSKVYVKHPADIPYKKLSFPEGFKWERVMNPFEDGCVTTQTDSLSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKHPADIPYKKLSFPEGFKWERVMNPFEDGCVTTQTDSLSLQDGCFTY 120  
OY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGIHKAALKDKDGGHYLVFKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGIHKAALKDKDGGHYLVFKSI 180

OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHLLFL 225  
DB 181 YMAKKPVQLPGYYVDADKLDITSHNEDYTYVEQYERTEGRHLLFL 225

RESULT 29  
AAE28919  
ID AAE28919 standard; protein; 225 AA.

AC AAE28919;

DT 27-DEC-2002 (first entry)

XX Discosoma sp. drFP583 (NFP-6) mutant protein, E5 (V105A) .

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW coloring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KW muteln.

OS Discosoma sp.  
OS Synthetic.

XX Key Location/Qualifiers  
FH Misc-difference 105 /note= "wild-type Val substituted with Ala"  
FT  
XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX 21-FEB-2001; 2001US-0270983P.

XX 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX WPI; 2002-691654/74.

XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.

PS Disclosure; Page; 80pp; English.

XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as coloring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification

XX Sequence 225 AA;

Query Match 99.7%; Score 1210; DB 5; Length 225;  
Best Local Similarity 99.6%; Pred. No. 2.9e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGLPFAWDI 60



Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBDDGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBDDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 30  
ABW00930  
ID ABW00930 standard; protein; 225 AA.  
XX  
AC ABW00930;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP), K83E.  
XX  
KW Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; mutant; mutein.  
XX  
OS Discosoma sp.  
OS Synthetic.

FH Key location/Qualifiers  
FT Misc-difference 83 /note= "Wild-type Lys is substituted with Glu"  
FT  
XX  
PN US2003170911-A1.  
XX  
PD 11-SEP-2003.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
PA (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
PI Tsien RY, Zacharias DA, Baird GS;  
XX  
DR WPI; 2003-802418/75.  
XX

PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.

XX Example 2; Page; 0pp; English.  
XX  
CC The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)

XX SQ Sequence 225 AA;  
Query Match 99.7%; Score 1210; DB 7; Length 225;  
. Best Local Similarity 99.6%; Pred. No. 2.9e-127;

Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBDDGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBDDGVVTVTQDSSLQDGCFTY 120  
QY 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 31  
AAG65510  
ID AAG65510 standard; protein; 226 AA.  
XX  
AC AAG65510;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Anthozoan red fluorescent protein synthetic mutant sequence.  
XX  
KW Fluorescent protein; Anthozoan; fluorescence; marker; FRET; mutant.  
XX  
OS Synthetic.  
XX  
PN WO200162919-A1.  
XX  
PD 30-AUG-2001.  
XX  
PF 13-FEB-2001; 2001WO-US004625.  
XX  
PR 23-FEB-2000; 2000US-0184732P.  
XX  
PA (AURO-) AURORA BIOSCIENCES CORP.  
XX  
PI Nelson D, Zamaira E, Tsien R;  
XX  
DR WPI; 2001-557704/62.  
DR N-PSDB; AAHA7656.  
XX

PT proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
PT functional red fluorescent proteins, and the encoding nucleic acids, with  
PT key mutations for improving the proteins function.

XX Disclosure; Page 87; 90pp; English.  
XX  
CC The invention provides a nucleic acid encoding functional red fluorescent  
CC CC protein (II) that differs from the sequence of an Anthozoan red  
CC fluorescent protein by at least one amino acid substitution, and with  
CC different fluorescent properties. The red fluorescent protein of the  
CC invention can be expressed by standard recombinant methodology. (II) are  
CC used a fluorescent markers and FRET partners. It is used for identifying  
CC protein-protein interactions. (II) is also suitable for multiplexed  
CC fluorescent analysis and FRET-based applications using existing Aegeorea  
CC fluorescent proteins. (II) has improved brightness, reduced spectral  
CC cross talk, and is rapidly and efficiently expressed in mammalian cells.  
CC The key mutations in the encoding nucleic acids provide improved folding,  
CC brightness, and create (II) with sharper, more defined excitation and  
CC emission peaks when expressed in mammalian cells. The present sequence  
CC represents the amino acid sequence of an improved synthetic mutant of an  
CC anthozoan fluorescent protein

XX SQ Sequence 226 AA;  
Query Match 99.7%; Score 1210; DB 4; Length 226;



Best Local Similarity 99.6%; Pred. No. 2.9e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 61  
QY 61 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
Db 62 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 121  
QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 122 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 181  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
Db 182 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 226

RESULT 32

ADE24109  
ID ADE24109 standard; protein; 242 AA.

AC ADE24109;  
XX  
DT 29-JAN-2004 (first entry)  
XX

DE Discosoma  $\delta$ p DsRed1 reporter protein.

KW site-specific DNA recombination; Cre recombinase;  
KW protein-transduction domain; gene activation; gene inactivation;  
KW chromosomal translocation; reporter protein.

XX Discosoma  $\delta$ p.

XX WO2003070931-A2.

XX PD 28-AUG-2003.

XX PF 19-FEB-2003; 2003WO-EP001680.

XX PR 21-FEB-2002; 2002DE-01007313.  
PR 16-JUL-2002; 2002DE-01032196.

XX PA (VISI-) VISION 7 GMBH.

XX PI Baum C, Will E, Osterag W, Klump H, Schiedlmeier B;

XX DR WPI; 2003-767353/72.

XX In vitro or in vivo site-specific DNA recombination, useful e.g. for gene  
PT inactivation, using Cre recombinase that lacks heterologous protein-  
PT transduction domain.

XX PS Claim 22; SEQ ID NO 6; 84pp; German.

CC This invention describes a novel method of site-specific DNA  
CC recombination in eukaryotic cells in vitro comprising using a  
CC bacteriophage P1 Cre recombinase that lacks heterologous protein-  
CC transduction domains. The invention also contains a reporter system for  
CC detecting site-specific DNA recombination in eukaryotic cells.  
CC Recombination with Cre recombinase lacking heterologous protein-  
CC transduction domains is used for therapeutic site-specific recombination  
CC in eukaryotic cells, in vivo or in vitro, e.g. for partial inactivation  
CC or activation of genes, also to create a chromosomal translocation. The  
CC modified recombinase can enter intact (not electrically or chemically  
CC treated) cells when added to culture medium and catalyzes recombination  
CC events in the nucleus with recombination rate over 50%, after a single  
CC application. The Cre-recombinase does not need to be supplied from  
CC nucleic acid introduced into the cell, so its effect is immediate but  
CC short-lived, limiting cytotoxic effects; eliminating the risk of non-  
CC specific integration of DNA into the genome and providing quick results.

CC This sequence represents the Discosoma  $\delta$ p reporter protein DsRed1 used in  
CC the method described in the invention.  
XX  
SQ Sequence 242 AA;

Query Match 99.7%; Score 1210; DB 7; Length 242;  
Best Local Similarity 99.6%; Pred. No. 3.2e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 15 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 74  
QY 61 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
Db 75 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTTQDSSLQDGCFTY 134  
QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180  
Db 135 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 194  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225  
Db 195 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 239

RESULT 33

ADL18132  
ID ADL18132 standard; protein; 545 AA.

XX ADL18132;

XX DT 06-MAY-2004 (first entry)

XX DE RFP:PS (N1a protease): AtoEP7:GFP fusion protein SEQ ID NO:52.

XX chimeric protein; signal protein; trafficking signal targeting;  
KW proteolytic cleavage site; protease; protease inhibitor.

XX Arabidopsis thaliana.

XX OS Tobacco vein mottling virus.

XX OS Synthetic.

XX PN WO2003014381-A1.

XX PD 20-FEB-2003.

XX PF 08-AUG-2002; 2002WO-KR001515.

XX PR 10-AUG-2001; 2001KR-00048123.

XX PA (AHRM-) AHRM BIOSYSTEMS INC.

XX PI Hwang I, Kim DH, Lee YU;

XX DR WPI; 2003-256596/25.

XX N-PSDB; ADL18131.

XX PT New chimeric protein, useful for detecting protease inhibitors inside the  
XX cell or tissue.

XX PS Example 2; SEQ ID NO 52; 214pp; English.

CC The present invention describes a chimeric protein comprising at least  
CC one signal protein that has a trafficking signal targeting to a  
CC subcellular organelle and at least one proteolytic cleavage site for a  
CC protease. The chimeric protein is constructed, so that: (a) the  
CC trafficking signals of all the signal proteins are inactivated by linking  
CC the proteolytic site or a signal masking protein through the proteolytic  
CC site to the N-or C-terminus of the signal proteins, and so the chimeric  
CC protein is present in cytosol; (b) the trafficking signal of at least one  
CC signal protein is activated when the proteolytic cleavage site is cleaved  
CC by the protease, and as a result at least one fragment protein that



CC includes the activated signal protein is a transported to a subcellular  
CC organelle; and (c) the chimeric protein is labelled with at least one  
CC fluorescent protein and the position and intensity distribution of the  
CC fluorescent label signal in the cell is altered depending on the cleavage  
CC by the protease. Also described: (1) a recombinant gene comprising a  
CC nucleic acid sequence encoding the chimeric protein which is constructed  
CC to express the chimeric protein in a cell; (2) a cell transformed with  
CC the recombinant gene or vector; (3) analysing the activity of a protease  
CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
CC detecting a protease inside a cell; (6) a nucleic acid comprising the  
CC sequence encoding the chimeric protein for detecting protease activity in  
CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
CC a protease inside a cell comprising the chimeric protein or the vector;  
CC (9) detecting a protease inside a cell or tissue; and (10) detecting a  
CC protease inhibitor in vivo. The chimeric protein is useful for detecting  
CC protease inhibitors inside the cell or tissue. The present sequence  
CC represents a fusion protein, which is used in the exemplification of the  
CC present invention.

XX Sequence 545 AA;

Query Match 99.7%; Score 1210; DB 7; Length 545;  
Best Local Similarity 99.6%; Pred. No. 1.1e-126;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 61  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSPEEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 120  
|||||  
DB 62 LSPQFQYGSKVYVKHPADIPDYKKLSPEEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 121  
  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
DB 122 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||  
DB 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 34  
ADL18156  
ID ADL18156 standard; protein; 548 AA.

XX AC ADL18156;  
XX DT 06-MAY-2004 (first entry)  
XX DE RFP:PS (HIV-1 protease):AtOEP7:GFP fusion protein SEQ ID NO:76.  
XX KW chimeric protein; signal protein; trafficking signal targeting;  
KW proteolytic cleavage site; protease; protease inhibitor.

XX Arabidopsis thaliana.  
OS Human immunodeficiency virus 1.  
OS Synthetic.

XX PN WO2003014381-A1.  
XX PD 20-FEB-2003.

XX PF 08-AUG-2002; 2002WO-KR001515.  
XX PR 10-AUG-2001; 2001KR-00048123.

XX PA (AHRA-) AHRAM BIOSYSTEMS INC.

XX PI Hwang I, Kim DH, Lee YJ;

XX DR WPI; 2003-256596/25.  
DR N-PSDB; ADL18155.

XX New chimeric protein, useful for detecting protease inhibitors inside the  
PT cell or tissue.

XX Example 3; SEQ ID NO 76; 214pp; English.

PS The present invention describes a chimeric protein comprising at least  
XX one signal protein that has a trafficking signal targeting to a  
CC subcellular organelle and at least one proteolytic cleavage site for a  
CC protease. The chimeric protein is constructed, so that: (a) the  
CC trafficking signals of all the signal proteins are inactivated by linking  
CC the proteolytic site or a signal masking protein through the proteolytic  
CC site to the N-or C- terminus of the signal proteins, and so the chimeric  
CC protein is present in cytosol; (b) the trafficking signal of at least one  
CC signal protein is activated when the proteolytic cleavage site is cleaved  
CC by the protease, and as a result at least one fragment protein that  
CC includes the activated signal protein is a transported to a subcellular  
CC organelle; and (c) the chimeric protein is labelled with at least one  
CC fluorescent protein and the position and intensity distribution of the  
CC fluorescent label signal in the cell is altered depending on the cleavage  
CC by the protease. Also described: (1) a recombinant gene comprising a  
CC nucleic acid sequence encoding the chimeric protein which is constructed  
CC to express the chimeric protein in a cell; (2) a cell transformed with  
CC the recombinant gene or vector; (3) analysing the activity of a protease  
CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
CC detecting a protease inside a cell; (6) a nucleic acid comprising the  
CC sequence encoding the chimeric protein for detecting protease activity in  
CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
CC a protease inside a cell comprising the chimeric protein or the vector;  
CC (9) detecting a protease inside a cell or tissue; and (10) detecting a  
CC protease inhibitor in vivo. The chimeric protein is useful for detecting  
CC protease inhibitors inside the cell or tissue. The present sequence  
CC represents a fusion protein, which is used in the exemplification of the  
CC present invention.

XX Sequence 548 AA;

Query Match 99.7%; Score 1210; DB 7; Length 548;  
Best Local Similarity 99.6%; Pred. No. 1.1e-126;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 61  
  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSPEEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 120  
|||||  
DB 62 LSPQFQYGSKVYVKHPADIPDYKKLSPEEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 121  
  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
DB 122 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||  
DB 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 35  
AAE28922  
ID AAE28922 standard; protein; 225 AA.

XX AAE28922;

XX AC AAE28922;  
XX DT 27-DEC-2002 (first entry)

XX DE Discosoma sp. drFP583 (NFP-6) mutant protein, E8 (N42H).

XX KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;



KW mutein.  
XX Discosoma sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 42 /note= "Wild-type Asn substituted with His"  
XX  
XX  
PN WO200268459-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 20-FEB-2002; 2002WO-US005749.  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S; Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX WPI; 2002-691654/74.  
XX  
XX  
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chlarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
PS Disclosure; Page; 80pp; English.  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as screens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. dFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. dFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX Sequence 225 AA;  
SQ

Query Match 99.6%; Score 1209; DB 5; Length 225;  
Best Local Similarity 99.6%; Pred. No. 3.8e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120

QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 36  
ABW00936  
ID ABW00936 standard; protein; 225 AA.  
XX  
AC ABW00936;  
XX

DT 15-JAN-2004 (first entry)  
XX  
XX Discosoma sp. red fluorescent protein (RFP), Y120H.  
DE  
XX  
XX Fluorescent protein; resonance energy transfer; pH; detection;  
KW red fluorescent protein; RFP; mutant; mutein.  
XX  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 120 /note= "Wild-type Tyr is substituted with His"  
FT  
XX  
XX  
PN US2003170911-A1.  
XX  
XX 11-SEP-2003.  
XX  
PD 26-FEB-2001; 2001US-00794308.  
XX  
PF 26-FEB-2001; 2001US-00794308.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
XX  
XX (TSIE/) TSIE R Y.  
PA (ZACH/) ZACHARIAS D A.  
PA (BAIR/) BAIRD G S.  
XX  
XX Tsien RY, Zacharias DA, Baird GS;  
PI  
XX  
XX WPI; 2003-802418/75.  
XX  
XX  
XX Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.  
XX  
XX  
PS Example 2; Page; 0pp; English.  
XX  
XX The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)  
XX  
XX Sequence 225 AA;  
SQ

Query Match 99.6%; Score 1209; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 3.8e-127;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60

QY 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGVTVTQDSSLQDGCFTY 120

QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 37  
ABW00931



ID	ABW00931	standard; protein; 225 AA.
XX		
AC	ABW00931;	
XX		
DT	15-JAN-2004	(first entry)
XX		
DE	Discosoma sp.	red fluorescent protein (RFP) , K83N.
XX		
KW	Fluorescent protein; resonance energy transfer; pH; detection;	
KW	red fluorescent protein; RFP; mutant; mutein.	
XX		
OS	Discosoma sp.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference 83	/note= "Wild-type Lys is substituted with Asn"
XX		
PN	US2003170911-A1.	
XX		
PD	11-SEP-2003.	
XX		
PF	26-FEB-2001; 2001US-00794308.	
XX		
PR	26-FEB-2001; 2001US-00794308.	
XX		
PA	(TSIE/) TSIE N Y.	
PA	(ZACH/) ZACHARIAS D A.	
PA	(BAIR/) BAIRD G S.	
XX		
PI	Tsien RY, Zacharias DA, Baird GS;	
XX		
DR	WPI; 2003-802418/75.	
XX		
PT	Fluorescent proteins containing a mutation that reduces or eliminates its ability to oligomerize which gives more reliable fluorescence resonance energy transfer results and are useful to detect molecule interaction,	
PT	enzymes, or sample pH.	
XX		
PS	Example 2; Page; 0pp; English.	
XX		
CC	The invention relates to a non-oligomerising fluorescent protein	
CC	containing a mutation that reduces or eliminates its ability to	
CC	oligomerise. The fluorescent protein gives more reliable fluorescence	
CC	resonance energy transfer results and are useful to detect molecule	
CC	interaction, enzymes, or sample pH. These are also used to identify	
CC	agents or conditions that regulate expression of control sequences. The	
CC	present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.	
CC	Note: This sequence is not shown in the specification, however this	
CC	sequence is constructed based on Discosoma sp. wild-type RFP protein	
CC	shown in page 30-31 (ABW00918)	
XX		
SQ	Sequence 225 AA;	
	Query Match	99.6%; Score 1209; DB 7; Length 225;
	Best Local Similarity	99.6%; Pred. No. 3.8e-127;
	Matches 224; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD1 60
DB	1	MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD1 60
QY	61	LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
DB	61	LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
QY	121	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
DB	121	KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
QY	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
DB	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

		RESULT 38
		ABB08835
ID	ABB08835	standard; protein; 226 AA.
XX		
AC	ABB08835;	
XX		
DT	29-MAY-2002	(first entry)
XX		
DE	Yeast optimised RFP SEQ ID NO 20.	
XX		
KW	Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;	
KW	Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;	
KW	Escherichia coli; green fluorescent protein; biotechnology.	
XX		
OS	Anthozoa.	
XX		
PN	DE20001395-U1.	
XX		
PD	15-MAR-2001.	
XX		
PF	27-JAN-2000; 2000DE-02001395.	
XX		
PR	27-JAN-2000; 2000DE-02001395.	
XX		
PA	(GPCB-) GPC BIOTECH AG.	
XX		
DR	WPI; 2002-228394/29.	
DR	N-PSDB; ABA95906.	
XX		
PT	New DNA encoding red fluorescent protein, useful as marker in	
PT	biotechnology, has sequence optimized for expression in eukaryotes,	
PT	especially yeast or plants.	
XX		
PS	Disclosure; Page 15; 19pp; German.	
XX		
CC	The invention relates to DNA (I) containing either sequence ABA95905 or	
CC	sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein	
CC	(yRFP). (I) are used to express red fluorescent protein (RFP) in	
CC	eukaryotes, especially yeast, especially Saccharomyces cerevisiae and	
CC	plants, especially dicotyledonous plants including Nicotiana tabacum or	
CC	Arabidopsis thaliana and also in prokaryotes, especially bacteria,	
CC	especially Escherichia coli. RFP is useful in the same way as green	
CC	fluorescent protein but is more generally applicable in modern	
CC	biotechnology. (I) are optimised for expression in yeast and so generate	
CC	RFP at higher levels with stronger fluorescence and thus lowers the	
CC	detection limit and gives a better signal-to-noise ratio. The present	
CC	sequence is that of the yeast optimised RFP	
XX		
SQ	Sequence 226 AA;	
	Query Match	99.6%; Score 1209; DB 5; Length 226;
	Best Local Similarity	100.0%; Pred. No. 3.8e-127;
	Matches 224; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	2	RSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD1L 61
DB	3	RSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD1L 62
QY	62	SPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTYK 121
DB	63	SPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTYK 122
QY	122	VKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSIY 181
DB	123	VKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSIY 182
QY	182	MAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
DB	183	MAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226



```
RESULT 39
ABW00932
ID ABW00932 standard; protein; 225 AA.
XX
XX AC ABW00932;
XX
XX 15-JAN-2004 (first entry)
XX
XX Discosoma sp. red fluorescent protein (RFP), K83P.
XX
XX Fluorescent protein; resonance energy transfer; pH; detection;
XX KW red fluorescent protein; RFP; mutant; mutein.
XX
XX Discosoma sp.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 83
FT /note= "Wild-type Lys is substituted with Pro"
XX
XX US2003170911-A1.
XX
XX 11-SEP-2003.
XX
XX 26-FEB-2001; 2001US-00794308.
XX
XX 26-FEB-2001; 2001US-00794308.
XX
XX (TSIE/) TSIE R Y.
XX PA (ZACH/) ZACHARIAS D A.
XX PA (BAIR/) BAIRD G S.
XX
XX Tsien RY, Zacharias DA, Baird GS;
XX WPI; 2003-802418/75.
XX
XX Fluorescent proteins containing a mutation that reduces or eliminates its
XX PT ability to oligomerize which gives more reliable fluorescence resonance
XX PT energy transfer results and are useful to detect molecule interaction,
XX PT enzymes, or sample pH.
XX
XX Example 2; Page; Opp; English.
XX
XX The invention relates to a non-oligomerizing fluorescent protein
XX CC containing a mutation that reduces or eliminates its ability to
XX CC oligomerize. The fluorescent protein gives more reliable fluorescence
XX CC resonance energy transfer results and are useful to detect molecule
XX CC interaction, enzymes, or sample pH. These are also used to identify
XX CC agents or conditions that regulate expression of control sequences. The
XX CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.
XX CC Note: This sequence is not shown in the specification, however this
XX CC sequence is constructed based on Discosoma sp. wild-type RFP protein
XX CC shown in page 30-31 (ABW00918)
XX
XX Sequence 225 AA;
SQ
Query Match 99.5%; Score 1208; DB 7; Length 225;
Best Local Similarity 99.6%; Pred. No. 4.9e-127;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
OY 61 LSPQFOYGSKVYVKGADIPDYKGLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120
Db 61 LSPQFOYGSKVYVKGADIPDYKGLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120
OY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
Db 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
OY 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTLVEQYERTEGRHHLEFL 225
```

```
Db 181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTLVEQYERTEGRHHLEFL 225
RESULT 40
ABW00935
ID ABW00935 standard; protein; 225 AA.
XX
XX AC ABW00935;
XX
XX 15-JAN-2004 (first entry)
XX
XX Discosoma sp. red fluorescent protein (RFP), K83M.
XX
XX Fluorescent protein; resonance energy transfer; pH; detection;
XX KW red fluorescent protein; RFP; mutant; mutein.
XX
XX Discosoma sp.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 83
FT /note= "Wild-type Lys is substituted with Met"
XX
XX US2003170911-A1.
XX
XX 11-SEP-2003.
XX
XX 26-FEB-2001; 2001US-00794308.
XX
XX 26-FEB-2001; 2001US-00794308.
XX
XX (TSIE/) TSIE R Y.
XX PA (ZACH/) ZACHARIAS D A.
XX PA (BAIR/) BAIRD G S.
XX
XX Tsien RY, Zacharias DA, Baird GS;
XX WPI; 2003-802418/75.
XX
XX Fluorescent proteins containing a mutation that reduces or eliminates its
XX PT ability to oligomerize which gives more reliable fluorescence resonance
XX PT energy transfer results and are useful to detect molecule interaction,
XX PT enzymes, or sample pH.
XX
XX Example 2; Page; Opp; English.
XX
XX The invention relates to a non-oligomerizing fluorescent protein
XX CC containing a mutation that reduces or eliminates its ability to
XX CC oligomerize. The fluorescent protein gives more reliable fluorescence
XX CC resonance energy transfer results and are useful to detect molecule
XX CC interaction, enzymes, or sample pH. These are also used to identify
XX CC agents or conditions that regulate expression of control sequences. The
XX CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.
XX CC Note: This sequence is not shown in the specification, however this
XX CC sequence is constructed based on Discosoma sp. wild-type RFP protein
XX CC shown in page 30-31 (ABW00918)
XX
XX Sequence 225 AA;
SQ
Query Match 99.5%; Score 1208; DB 7; Length 225;
Best Local Similarity 99.6%; Pred. No. 4.9e-127;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
OY 61 LSPQFOYGSKVYVKGADIPDYKGLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120
Db 61 LSPQFOYGSKVYVKGADIPDYKGLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120
OY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
```



```

Db      121  KYKFIGVNFPSDGPVMQKKTGWBASTERLYPRDGLKGEIHKALKDKGHYLVEFKSI 180
QY      181  YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      181  YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 41
ABW00939 ID ABW00939 standard; protein; 225 AA.
XX AC ABW00939;
XX DT 15-JAN-2004 (first entry)
XX DE Discosoma sp. red fluorescent protein (RFP), K70M.
XX KW Fluorescent protein; resonance energy transfer; pH; detection;
XX KW red fluorescent protein; RFP; mutant; muteln.
XX OS Discosoma sp.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 70 /note= "Wild-type Lys is substituted with Met"
XX PN US2003170911-A1.
XX PD 11-SEP-2003.
XX PF 26-FEB-2001; 2001US-00794308.
XX PR 26-FEB-2001; 2001US-00794308.
XX TS (TSIE/) TSIE R Y.
XX PA (ZACH/) ZACHARIAS D A.
XX PA (BAIR/) BAIRD G S.
XX PI Telen RY, Zacharias DA, Baird GS;
XX DR WPI; 2003-802418/75.
XX PT Fluorescent proteins containing a mutation that reduces or eliminates its
XX PT ability to oligomerize which gives more reliable fluorescence resonance
XX PT energy transfer results and are useful to detect molecule interaction,
XX PT enzymes, or sample pH.
XX PS Example 2; Page; 0pp; English.
XX CC The invention relates to a non-oligomerising fluorescent protein
XX CC containing a mutation that reduces or eliminates its ability to
XX CC oligomerise. The fluorescent protein gives more reliable fluorescence
XX CC resonance energy transfer results and are useful to detect molecule
XX CC interaction, enzymes, or sample pH. These are also used to identify
XX CC agents or conditions that regulate expression of control sequences. The
XX CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.
XX CC Note: This sequence is not shown in the specification, however this
XX CC sequence is constructed based on Discosoma sp. wild-type RFP protein
XX CC shown in page 30-31 (ABW00918)
XX SQ Sequence 225 AA;

Query Match 99.5%; Score 1208; DB 7; Length 225;
Best Local Similarity 99.6%; Pred. No. 4.9e-127;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY      1 MRSSKNVIKEFMRFKVRMEGTIVNGHFEIEGEGGRPYEGHNTVYKLKVTGKGPLPFAMDI 60
Db      1 MRSSKNVIKEFMRFKVRMEGTIVNGHFEIEGEGGRPYEGHNTVYKLKVTGKGPLPFAMDI 60
QY      61 LSPQFOYGSKVYVKAHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTITQDSSLQDGCFTY 120

```

Db	61	LSPOFGYSMYVYKHPADI PDYKCLSPFEGFKMERVMNFEDGGVVTQTDDSLQDGCFTY	120
QY	121	KVKFIGVNFPSDGPVMQKKTMGWEASTELYP RDGLKGEIHKALKLDGGHYLVFPKSI	180
Db	121	KVKFIGVNFPSDGPVMQKKTMGWEASTELYP RDGLKGEIHKALKLDGGHYLVFPKSI	180
QY	181	YMAKKPVLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225
Db	181	YMAKKPVLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225
RESULT 42			
ADH34500	ID	ADH34500 standard; protein; 225 AA.	
XX	AC	ADH34500;	
XX	DT	11-MAR-2004 (first entry)	
XX	DE	Discosoma sp. DsRed mutant S148A/S203A.	
KW	KW	Chromoprotein; fluorescent protein; CP, FP; interconverted mutant;	
KW	KW	Cnidarian; Anthozoan; labelling; colouring agents; pigment;	
KW	KW	analyte detection assay; selectable marker; sunscreen; selective filter;	
KW	KW	fluorescence resonance energy transfer; FRET; biosensor;	
KW	KW	whole cell marker; second messenger detector; in vivo marker;	
KW	KW	fluorescence activated cell sorting; fluorescent timer;	
KW	KW	red fluorescent protein; DsRed; mutant; muteln.	
XX	OS	Synthetic.	
OS	OS	Discosoma sp.	
XX	XX	Key	Location/Qualifiers
FT	FT	Misc-difference 146	/note= "Ala replaces wild-type Ser. Corresponds to GFP residue 148"
FT	FT	Misc-difference 197	/note= "Ala replaces wild-type Ser. Corresponds to GFP residue 203"
XX	PN	WO2003057833-A2.	
XX	PD	17-JUL-2003.	
XX	PF	23-DEC-2002; 2002MO-US041418.	
XX	PR	26-DEC-2001; 2001US-0343128P.	
PA	PA	(CLON-) CLONTECH LAB INC.	
PI	PI	Bulina ME, Chudakov D, Lukyanov KA,	
XX	DR	WPI; 2003-607998/57.	
PT	PT	Novel nucleic acid encoding interconverted mutant of chromo-or	
PT	PT	fluorescent protein which are useful as biosensors, coloring agents.	
XX	PS	Example 1; Page; 56pp; English.	
XX	XX	The invention relates to interconverted mutants of chromoproteins (CP) or	
CC	CC	fluorescent proteins (FP) and nucleic acids encoding them. The mutant is	
CC	CC	derived from a Cnidarian species, preferably a non-bioluminescent	
CC	CC	Cnidarian species, and most preferably an Anthozoan species. The	
CC	CC	invention is based on the finding that although green fluorescent protein	
CC	CC	(GFP)-like chromoproteins and fluorescent proteins exhibit some degree of	
CC	CC	homology, there are certain positions (referred to as 148, 165, 167 and	
CC	CC	203; numbering corresponds to GFP) that are occupied by noticeably	
CC	CC	different residues in the two types of proteins. Mutagenesis of the	
CC	CC	residues in these key positions in, for example, a fluorescent protein,	
CC	CC	to those found in a chromoprotein is therefore proposed to confer	
CC	CC	chromoprotein activity on the fluorescent protein mutant, with	
CC	CC	chromoproteins being able to be converted into fluorescent proteins in a	



CC similar manner. The invention also relates to expression constructs, vectors, host cells and host cell progeny comprising a nucleic acid of the invention; the recombinant production of an interconverted chromoprotein or fluorescent protein mutant; and antibodies specific for interconverted mutant proteins of the invention. The interconverted mutants are useful in any application that employs a chromoprotein or fluorescent protein. Fluorescent protein mutants having chromoprotein activity can be useful as colouring agents in, for example, food compositions, pharmaceuticals, cosmetics and living organisms. Proteins with chromoprotein activity are also useful as labels in biological analyte detection assays, as selectable markers in recombinant DNA applications (e.g. the production of transgenic cells and organisms), and are also useful as sunscreens and selective filters. Chromoprotein mutants having fluorescent protein activity useful in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, as markers of whole cells to detect changes in multicellular reorganisation and migration, as second messenger detectors, as in vivo markers in animals (e.g., transgenic animals), in fluorescence activated cell sorting applications, in protease cleavage assays, and in assays to determine the phospholipid composition in biological membranes. Proteins with fluorescent protein activity can also be used as fluorescent timers, where the switch of one fluorescent colour to another (e.g., green to red) is concomitant with the ageing of the protein and is useful for determination of the activation or deactivation of gene expression. The present sequence represents a *DiscoSoma* sp. red fluorescent protein DsRed mutant generated in an example of the invention. The present sequence is not shown in the specification, but was derived from the wild-type DsRed sequence (ADH34489) shown in Fig 1 and the information provided on page 42.

CC XX Sequence 225 AA;

Query Match 99.5%; Score 1208; DB 7; Length 225;  
Best Local Similarity 99.1%; Pred. No. 4.9e-127;  
Matches 223; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 61 LSPQFOYGSKVYVKHPADIPDYKGLSPPEGFKWERVMNFEEDGVATVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKGLSPPEGFKWERVMNFEEDGVATVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFPKSI 180  
121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFPKSI 180  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225  
181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225  
Db 181 YMAKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225

RESULT 43  
AAE28921  
ID AAE28921 standard; protein; 225 AA.  
XX AAE28921;  
AC AAE28921;  
XX 27-DEC-2002 (first entry)  
DT 27-DEC-2002 (first entry)  
XX  
DE *DiscoSoma* sp. drFP583 (NFP-6) mutant protein, E5 (V105A+S197T).

XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KM fluorescence activated cell sorting application; fluorescent timer;  
KM biosensor; fluorescence resonance energy transfer application; FRET;  
KM colouring agent; recombinant DNA application; analyte detection assay;  
KM sunsreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KM muteln.

XX *DiscoSoma* sp.  
OS Synthetic.  
XX

FH Key Location/Qualifiers

FT Misc-difference 105 /note= "Wild-type Val substituted with Ala"

FT Misc-difference 217 /note= "Wild-type Ser substituted with Thr"

PN WO200268459-A2.

PD 06-SEP-2002.

PF 20-FEB-2002; 2002WO-US005749.

PR 21-FEB-2001; 2001US-0270983P.

PR 04-DEC-2001; 2001US-00006922.

PA (CLON-) CLONTECH LAB INC.

PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;

DR WPI; 2002-691654/74.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant of an aggregating Cnidarian chromo- or fluorescent protein or mutant for analyte detection assays or fluorescence activated cell sorting applications.

PS Disclosure; Page; 80pp; English.

CC The invention relates to nucleic acid molecules encoding non-aggregating chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are useful in analyte detection assays, as colouring agents, as markers in recombinant DNA applications, as sunsreens or filters, in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in screening assays, as second messenger detectors, in fluorescence activated cell sorting applications, in protease cleavage assays or as fluorescent timers. The present sequence is *DiscoSoma* sp. drFP583 (NFP-6) mutant protein of the invention. Note: This sequence is not shown in the specification, but is derived from *DiscoSoma* sp. drFP583 (NFP-6) wild-type protein shown as SEQ ID NO:8 (AAE28833) in page 70-71 of the specification

XX Sequence 225 AA;

Query Match 99.4%; Score 1207; DB 5; Length 225;  
Best Local Similarity 99.1%; Pred. No. 6.3e-127;  
Matches 223; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
1 MRSSKNVKEFMRFKVMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 61 LSPQFOYGSKVYVKHPADIPDYKGLSPPEGFKWERVMNFEEDGVATVTQDSSLQDGCFTY 120  
61 LSPQFOYGSKVYVKHPADIPDYKGLSPPEGFKWERVMNFEEDGVATVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFPKSI 180  
121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFPKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFPKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225  
181 YMAKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225  
Db 181 YMAKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERTEGRHHLLFL 225

RESULT 44  
AAE17541  
ID AAE17541 standard; protein; 225 AA.  
XX AAE17541;  
AC AAE17541;  
XX 22-APR-2002 (first entry)  
DT 22-APR-2002 (first entry)  
XX

XX *DiscoSoma* sp. fluorescent protein E5.



KW	Fluorescent timer protein; protein movement; translocation; trafficking;
KM	promoter activity; gene expression; transgenic plant; gene modification;
KM	protein age; E5.
XX	
OS	DiscoSoma sp.
XX	
FH	Key
FT	Misc-difference 105
FT	/note= "Wild type Val substituted with Ala"
FT	Misc-difference 197
FT	/note= "Wild type Ser substituted with Thr"
XX	
PN	WO200196373-A2.
XX	
PD	20-DEC-2001.
XX	
PF	13-JUN-2001; 2001WO-US019097.
XX	
PR	14-JUN-2000; 2000US-0211607P.
XX	
PA	(CLON-) CLONTECH LAB INC.
XX	
PI	Pradkov AF, Terakikh A;
XX	
DR	WPI; 2002-154595/20.
XX	
PS	N-PSDB; AAD28208.
XX	
PT	Claim 5; Fig 2; 89pp; English.
XX	
CC	The invention relates to a fluorescent timer protein having an emission spectrum that changes over time after synthesis from a first wavelength to a second wavelength. The fluorescent timer proteins are useful in monitoring the activity of a promoter, determining the age of a protein, identifying an agent that modulates the activity of a promoter and in enriching a population of cells comprising a fluorescent timer protein. The fluorescent timer proteins are also useful for assessing gene expression during development of a multicellular organism or during cellular differentiation, in response to a drug or other inducer of promoter activity, as a reporter to serve as a read-out of promoter activity, monitoring intracellular protein movement or translocation, protein trafficking, or protein stability, to investigate temporal aspects of the activity of a regulatory element, for determining cell fate during development and organ remodelling, in spatial and temporal visualisation of newly synthesised proteins and accumulated proteins, and in distinguishing between newly formed and pre-existing structures, e.g. membrane junctions and extracellular matrix components. The fluorescent timer proteins may further be used to investigate where photobleaching techniques are employed, as detectable labels, as selectable markers, as biosensors in prokaryotic and eukaryotic cells, in protease cleavage assays, and as second messenger detectors. The nucleic acids can be used to generate transgenic, non-human plants or animals or site-specific gene modifications in cell lines. The present sequence is DiscoSoma sp. E5
CC	fluorescent protein derived from humanised wild-type Anthozoa protein dFP583 by substituting Val to Ala at 105 and Ser to Thr at 197
CC	
XX	
XX	
SO	Sequence 225 AA;
QY	Query Match
	Best local Similarity 99.4%; Score 1207; DB 5; Length 225;
	Matches 223; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVKTKGGPLPFAMDI 60
DB	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVKTKGGPLPFAMDI 60
QY	61 LSPQFGYSKVYVVKHPADIPDYKKLSFPEGFKMERVWNFEDGVVTVTQDSSLQDGCFTY 120

Query	Match	Score	1207;	DB	5;	Length	225;
Db	61 LSPQFYGSKVYVGHADIPDYKKLSFPEGFKMERVVMNFEDGSVAITVTQDSSLQDGCPIY	120					
Qy	121 KVKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGETHKALKLDKGHYLVEFKSI	180					
Db	121 KVKFIGVNFPSDGPVWQKKTMGWEASTERLYPRDGVLKGETHKALKLDKGHYLVEFKSI	180					
Qy	181 YMAKPPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225						
Db	181 YMAKPPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHFL 225						
RESULT 45							
ID	ADY51734 standard; protein; 225 AA.						
XX	ADY51734;						
XX	05-MAY-2005 (first entry)						
XX	Discosoma sp. red fluorescent protein (RFP) mutant I125R.						
DE	fluorescence; mutagenesis; red fluorescent protein; protein interaction;						
KW	muteln.						
XX	Discosoma sp.						
OS	Synthetic.						
XX	Key	Location/Qualifiers					
FH	Misc-difference 125	/note= "Wild type Ile substituted by Arg"					
FT							
XX	WO200268605-A2.						
PN							
XX	06-SEP-2002.						
PD							
XX	26-FEB-2002; 2002WO-US006063.						
PF							
XX	26-FEB-2001; 2001US-00794308.						
PR	24-MAY-2001; 2001US-00866538.						
XX							
PA	(REGC ) UNIV CALIFORNIA.						
XX							
PI	Tsien RY, Baird GS, Campbell RE, Zacharias DA;						
XX							
DR	WPI, 2002-713372/77.						
XX							
PT	New non-oligomerizing fluorescent protein containing at least one						
PT	mutation that reduces or eliminates the ability of the protein to						
PT	oligomerize, useful for making better and new assays for molecular						
PT	biology.						
XX							
PS	Claim 11; Page; 117pp; English.						
CC	This invention relates to a novel non-oligomerizing fluorescent protein.						
CC	Specifically, it refers to the presence of at least one mutation in the						
CC	fluorescent protein that reduces or eliminates the ability of the protein						
CC	to oligomerize. The present invention describes fluorescent proteins						
CC	derived from naturally occurring green or red fluorescent proteins and						
CC	provides a fusion protein that comprises a non-oligomerizing fluorescent						
CC	protein linked to at least one protein of interest. As such, these fusion						
CC	proteins can be used in methods and compositions to determine the pH of a						
CC	sample, or whether the sample contains an enzyme, molecule or agent that						
CC	regulates the activity of an expression control sequence. Furthermore,						
CC	they may be used to identify a specific interaction of molecules, such						
CC	that they are useful for improving or developing new assays in the field						
CC	of molecular biology. This polypeptide sequence is the field						
CC	fluorescent protein (RFP) mutant of the invention. NOTE: This sequence is						
CC	not given in the invention but it is derived from Seqid 12 and						
CC	information provided in the claims.						
XX							
XX	Sequence 225 AA;						
XX							



Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAWDI 60  
QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFPKSI 180  
Db 121 KVKFRGVNFPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFPKSI 180  
QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 46  
ABW00941 ID ABW00941 standard; protein; 225 AA.

XX AC ABW00941;  
XX DT 15-JAN-2004 (first entry)

XX DE Discosoma sp. red fluorescent protein (RFP), I125R.

KW Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; muteln.

XX OS Discosoma sp.  
OS Synthetic.

XX FH Key Location/Qualifiers  
PT Misc-difference 125 /note= "wild-type Ile is substituted with Arg"

XX PN US2003170911-A1.

XX PD 11-SEP-2003.

XX PF 26-FEB-2001; 2001US-00794308.

XX PR 26-FEB-2001; 2001US-00794308.

XX PA (TSIE/) TSIENT R Y.

XX PA (ZACH/) ZACHARIAS D A.

XX PA (BAIR/) BAIRD G S.

XX PI Tsien RY, Zacharias DA, Baird GS;

XX DR WPI; 2003-802418/75.

XX PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.

XX PS Example 3; Page; 0pp; English.

XX CC The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The  
CC present sequence is Discosoma sp. red fluorescent protein (RFP) mutant.  
CC Note: This sequence is not shown in the specification, however this  
CC sequence is constructed based on Discosoma sp. wild-type RFP protein  
CC shown in page 30-31 (ABW00918)

XX SQ Sequence 225 AA;

Query Match 99.4%; Score 1207; DB 7; Length 225;  
Best Local Similarity 99.6%; Pred. No. 6.3e-127;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVCLKVTKGGLPFAWDI 60

QY 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFPKSI 180  
Db 121 KVKFRGVNFPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFPKSI 180

QY 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 47  
ABW00940 ID ABW00940 standard; protein; 225 AA.

XX AC ABW00940;

XX DT 15-JAN-2004 (first entry)

XX DE Discosoma sp. red fluorescent protein (RFP), I125K.

KW Fluorescent protein; resonance energy transfer; pH; detection;  
KM red fluorescent protein; RFP; mutant; muteln.

XX OS Discosoma sp.  
OS Synthetic.

XX FH Key Location/Qualifiers  
PT Misc-difference 125 /note= "wild-type Ile is substituted with Lys"

XX PN US2003170911-A1.

XX PD 11-SEP-2003.

XX PF 26-FEB-2001; 2001US-00794308.

XX PR 26-FEB-2001; 2001US-00794308.

XX PA (TSIE/) TSIENT R Y.

XX PA (ZACH/) ZACHARIAS D A.

XX PA (BAIR/) BAIRD G S.

XX PI Tsien RY, Zacharias DA, Baird GS;

XX DR WPI; 2003-802418/75.

XX PT Fluorescent proteins containing a mutation that reduces or eliminates its  
PT ability to oligomerize which gives more reliable fluorescence resonance  
PT energy transfer results and are useful to detect molecule interaction,  
PT enzymes, or sample pH.

XX PS Example 3; Page; 0pp; English.

XX CC The invention relates to a non-oligomerising fluorescent protein  
CC containing a mutation that reduces or eliminates its ability to  
CC oligomerise. The fluorescent protein gives more reliable fluorescence  
CC resonance energy transfer results and are useful to detect molecule  
CC interaction, enzymes, or sample pH. These are also used to identify  
CC agents or conditions that regulate expression of control sequences. The











XX (CLON-) CLONTECH LAB INC.  
PA Bulina ME, Chudakov D, Lukyanov KA;  
XX WPI; 2003-607998/57.  
DR  
XX Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
XX Example 1; Page; 56pp; English.  
XX  
CC The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs,  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can useful as colouring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and  
CC are also useful as sunscreens and selective filters. Chromoprotein  
CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein Dared mutant generated  
CC in an example of the invention. The present sequence is not shown in the  
CC specification, but was derived from the wild-type Dared sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.  
XX  
SQ Sequence 225 AA;

Query Match 99.3%; Score 1205; DB 7; Length 225;  
Best Local Similarity 99.1%; Pred. No. 1.1e-126;  
Matches 223; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MRSKVNLIKFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
Db 1 MRSKVNLIKFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFGSKVYVKGHPADIPDYKKLSPEEGFKMERVMNFBDDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGSKVYVKGHPADIPDYKKLSPEEGFKMERVMNFBDDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFTGVNFPSPDGPVMQKTMGWEASTERTLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

OY 181 YMAKPVQLPGYYVVDISKDITSHNEDYTIIVEQYERTGRHLLFL 225  
Db 181 YMAKPVQLPGYYVVDISKDITSHNEDYTIIVEQYERTGRHLLFL 225  
RESULT 52  
ABP56678  
ID ABP56678 standard; protein; 240 AA.  
XX  
AC ABP56678;  
XX  
DT 25-MAR-2003 (first entry)  
XX  
DE Mammalian codon optimised Discosoma red fluorescent protein.  
XX  
KW Mammalian codon optimised Discosoma red fluorescent protein; Discosoma;  
KW red fluorescent protein; directed evolution.  
XX  
OS Discosoma sp.  
OS Mammalia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 24 /note= "possible mutations at this point are N24S, N24R  
FT and N24H"  
FT Misc-difference 125 /note= "possible mutations at this point are F125L and  
FT F125V"  
FT Misc-difference 164 /note= "a possible mutation at this point is K164M"  
FT Misc-difference 183 /note= "a possible mutation at this point is M183K"  
FT Misc-difference 206 /note= "encoded by AAC"  
FT  
XX  
PN WO200294992-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 20-MAY-2002; 2002WO-US015968.  
XX  
PR 18-MAY-2001; 2001US-0291871P.  
XX  
PA (RIGB-) RIGEL PHARM INC.  
XX  
PI Peelle B;  
XX  
XX WPI; 2003-120798/11.  
DR N-PSDB; AB222476.  
DR  
XX  
XX  
PT New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.  
XX  
XX  
PS Claim 1; Fig 1; 22pp; English.  
XX  
XX The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising the  
CC an isolated nucleic acid encoding (I); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC characterisation and localisation (e.g. as a fusion protein), for target  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution



CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein, which is used in an  
CC example from the present invention

XX Sequence 240 AA;

Query Match 99.3%; Score 1205; DB 6; Length 240;  
Best Local Similarity 99.1%; Pred. NO. 1.2e-126;  
Matches 223; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVTKGKGPPLPFAMDI 60  
:|||||  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVTKGKGPPLPFAMDI 61  
61 LSPQFOYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
|||||  
Db 62 LSPQFOYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 121  
121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
|||||  
Db 122 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 181  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 182 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 226

RESULT 53  
AAE28923  
ID AAE28923 standard; protein; 225 AA.

XX AC AAE28923;  
DT 27-DEC-2002 (first entry)  
DE Discosoma sp. drFP583 (NFP-6) mutant protein, E83 (N42H+V71A+I180H).  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KM muteln.

XX Discosoma sp.  
OS Synthetic.

XX Key Location/Qualifiers  
FH Misc-difference 42 /note= "Wild-type Asn substituted with His"  
FT Misc-difference 71 /note= "Wild-type Val substituted with Ala"  
FT Misc-difference 180 /note= "Wild-type Ile substituted with Val"  
FT /note= "Wild-type Ser substituted with Thr"

XX WO200268459-A2.

XX 06-SEP-2002.

XX 20-FEB-2002; 2002WO-US005749.

XX 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;

XX WPI; 2002-691654/74.

PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.

XX Disclosure; Page; 80pp; English.

CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification

XX Sequence 225 AA;

Query Match 99.2%; Score 1204; DB 5; Length 225;  
Best Local Similarity 98.7%; Pred. NO. 1.4e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVTKGKGPPLPFAMDI 60  
|||||  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLVTKGKGPPLPFAMDI 60  
61 LSPQFOYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
|||||  
Db 61 LSPQFOYGSKVVYKHADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGVLKGEIHKALKKDGCHYLVEFKSV 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225

RESULT 54  
AAE28925  
ID AAE28925 standard; protein; 225 AA.

XX AAE28925;

DT 27-DEC-2002 (first entry)

DE Discosoma sp. drFP583 (NFP-6) mutant protein, AG4 (V71M+V105A+S197T).  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KM muteln.

XX Discosoma sp.  
OS Synthetic.

XX Key Location/Qualifiers  
FH Misc-difference 71 /note= "Wild-type Val substituted with Met"  
FT Misc-difference 105 /note= "Wild-type Val substituted with Ala"  
FT Misc-difference 197 /note= "Wild-type Ser substituted with Thr"  
FT /note= "Wild-type Ser substituted with Thr"

XX WO200268459-A2.



PD 06-SEP-2002.  
XX  
XX 20-FEB-2002; 2002WO-US005749.  
XX  
XX 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX  
XX WPI; 2002-691654/74.  
XX  
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX Disclosure; Page; 80pp; English.  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX Sequence 225 AA;  
SQ

Query Match 99.2%; Score 1204; DB 5; Length 225;  
Best Local Similarity 98.7%; Pred. No. 1.4e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWBRVMNFEDEGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWBRVMNFEDEGVATVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVVVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 55  
AAE28924  
ID AAE28924 standard; protein; 225 AA.  
XX  
XX AAE28924;  
XX  
XX 27-DEC-2002 (first entry)  
XX  
XX Discosoma sp. drFP583 (NFP-6) mutant protein, E57 (V105A+I161T+S197A).  
XX  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KM fluorescence activated cell sorting application; fluorescent timer;  
KM biosensor; fluorescence resonance energy transfer application; FRET;  
KM colouring agent; recombinant DNA application; analyte detection assay;  
KM sunsreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
XX  
XX

OS Discosoma sp.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 105  
FT /note= "Wild-type Val substituted with Ala"  
FT Misc-difference 161  
FT /note= "Wild-type Ile substituted with Thr"  
FT Misc-difference 197  
FT /note= "Wild-type Ser substituted with Ala"  
XX  
XX WO200268459-A2.  
XX  
XX 06-SEP-2002.  
XX  
XX 20-FEB-2002; 2002WO-US005749.  
XX  
XX 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX  
XX WPI; 2002-691654/74.  
XX  
XX New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX Disclosure; Page; 80pp; English.  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX Sequence 225 AA;  
SQ

Query Match 99.0%; Score 1202; DB 5; Length 225;  
Best Local Similarity 98.7%; Pred. No. 2.3e-126;  
Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWBRVMNFEDEGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKWBRVMNFEDEGVATVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVVVDAKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 56  
ADH34503  
ID ADH34503 standard; protein; 225 AA.  
XX  
XX



AC ADH34503;  
XX 11-MAR-2004 (first entry)  
XX Discosoma sp. DsRed mutant S148A/I165S/S203A.  
DE  
XX Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed; mutant; mutein.  
XX  
OS Synthetic.  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 148"  
FT Misc-difference 161 /note= "Ser replaces wild-type Ile. Corresponds to GFP  
FT residue 165"  
FT Misc-difference 197 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 203"  
XX  
XX WO2003057833-A2.  
PN  
PD 17-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US041418.  
PF  
XX 26-DEC-2001; 2001US-0343128P.  
PR  
XX (CLON-) CLONTECH LAB INC.  
PA  
XX Bulina ME, Chudakov D, Lukyanov KA;  
PI  
XX WPI; 2003-607998/57.  
DR  
XX Novel nucleic acid encoding interconverted mutant of chromo-or  
PT fluorescent protein which are useful as biosensors, coloring agents.  
XX  
XX Example 1; Page; 56pp; English.  
PS  
XX The invention relates to interconverted mutants of chromoproteins (CP) or  
CC fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
CC derived from a Cnidarian species, preferably a non-bioluminescent  
CC Cnidarian species, and most preferably an Anthozoan species. The  
CC invention is based on the finding that although green fluorescent protein  
CC (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
CC homology, there are certain positions (referred to as 148, 165, 167 and  
CC 203; numbering corresponds to GFP) that are occupied by noticeably  
CC different residues in the two types of proteins. Mutagenesis of the  
CC residues in these key positions in, for example, a fluorescent protein,  
CC to those found in a chromoprotein is therefore proposed to confer  
CC chromoprotein activity on the fluorescent protein mutant, with  
CC chromoproteins being able to be converted into fluorescent proteins in a  
CC similar manner. The invention also relates to expression constructs, with  
CC vectors, host cells and host cell progeny comprising a nucleic acid of  
CC the invention; the recombinant production of an interconverted  
CC chromoprotein or fluorescent protein mutant; and antibodies specific for  
CC interconverted mutant proteins of the invention. The interconverted  
CC mutants are useful in any application that employs a chromoprotein or  
CC fluorescent protein. Fluorescent protein mutants having chromoprotein  
CC activity can be useful as colouring agents in, for example, food  
CC compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
CC with chromoprotein activity are also useful as labels in biological  
CC analyte detection assays, as selectable markers in recombinant DNA  
CC applications (e.g. the production of transgenic cells and organisms), and  
CC are also useful as sunscreens and selective filters. Chromoprotein

CC mutants having fluorescent protein activity useful in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, as markers of whole cells to detect  
CC changes in multicellular reorganisation and migration, as second  
CC messenger detectors, as in vivo markers in animals (e.g., transgenic  
CC animals), in fluorescence activated cell sorting applications, in  
CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing or deactivation of gene expression. The present sequence  
CC represents a Discosoma sp. red fluorescent protein DsRed mutant generated  
CC in an example of the invention. The present protein DsRed mutant generated  
CC specification, but was derived from the wild-type DsRed sequence  
CC (ADH34489) shown in Fig 1 and the information provided on page 42.  
XX  
SQ Sequence 225 AA;  
XX  
XX Query Match 99.0%; Score 1202; DB 7; Length 225;  
XX Best Local Similarity 98.7%; Pred. No. 2.3e-126;  
XX Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MRSSKNVTKFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
DB 1 MRSSKNVTKFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKGHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKVYVKGHPADIPDYKKLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNPSSDGPVMQKTMGWEASTERLYPRDGLVKGFIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNPSSDGPVMQKTMGWEASTERLYPRDGLVKGESHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYVVDADKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYVVDADKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
XX  
XX RESULT 57  
XX ADH34502  
XX ID ADH34502 standard; protein; 225 AA.  
XX  
XX AC ADH34502;  
XX  
XX 11-MAR-2004 (first entry)  
XX  
XX Discosoma sp. DsRed mutant S148A/K167M/S203A.  
DE  
XX Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed; mutant; mutein.  
XX  
XX Synthetic.  
OS Discosoma sp.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 146 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 148"  
FT Misc-difference 163 /note= "Met replaces wild-type Lys. Corresponds to GFP  
FT residue 167"  
FT Misc-difference 197 /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 203"  
XX  
XX WO2003057833-A2.







Query Match 99.0%; Score 1202; DB 6; Length 240;  
Best Local Similarity 98.7%; Pred. No. 2.5e-126;  
Matches 222; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD 60  
:|||||  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD 61

QY 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 120  
|||||  
Db 62 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 121

QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
Db 122 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||  
Db 182 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 59  
ABP70039  
ID ABP70039 standard; protein; 225 AA.  
XX AC ABP70039;  
XX DT 06-AUG-2003 (revised)  
DT 22-JAN-2003 (first entry)  
XX DE Colour facilitating molecule (CFM) related sequence #SEQ ID 241.

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;  
KM chromophore; biomatix; transgenic animal; colouring agent;  
KW flower industry; expression marker; reporter molecule; photon trap;  
KW UV sink; sunscreen.

XX Unidentified.  
OS  
XX WO200270703-A2.  
PN 12-SEP-2002.  
XX PD 01-MAR-2002; 2002WO-GB000928.  
XX PF 02-MAR-2001; 2001US-0273227P.  
XX PR 21-MAR-2001; 2001AU-00003874.  
PR 15-OCT-2001; 2001US-0329816P.  
XX (NUFA-) NUFARM LTD.  
PA (UYOU ) UNIV QUEENSLAND.  
PA (JONE/) JONES E L.  
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;  
PI Hoegh-Guldberg IO, Prescott M;  
PI WPI; 2002-740765/80.  
XX DR  
XX Novel color-facilitating molecule for producing a biomatix, has a  
PT polypeptide which alone/along with molecules imparts altered visual  
PT characteristics to cells in the absence of excitation by extraneous non-  
PT white light.  
XX  
XX Example 19; Page 499; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)  
CC comprising a polypeptide which, in a cell, alone or together with one or  
CC more other molecules imparts an altered visual characteristic to the cell  
CC when visualised by a human eye in the absence of excitation by extraneous  
CC non-white light or particle emission. CFMs are useful for producing a  
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or  
CC red coloured fleece. They are useful for producing coloured plant

CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other  
CC uses include transducing or intensifying an image, providing additional  
CC light for growing phototropic organisms e.g. algae and/or corals, for  
CC coating materials that experience UV damage e.g. plastics and car  
CC upholstery. CFMs are useful in the flower industry, in the development of  
CC new varieties of flowering plants. Other contemplated uses include,  
CC expression markers, general reporter molecules, photon traps, UV sinks or  
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental  
CC fungal species, and in fruits and vegetables to enhance their  
CC marketability. CFMs embedded in a gel matrix improve image quality in  
CC situations of distorted light spectra (biomatix). The first all-protein  
CC chromophore to be isolated was Green Fluorescent protein (GFP). The  
CC sequences given in records ABP69924-ABP70048 represent CFM related amino  
CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
SQ Sequence 225 AA;

Query Match 98.8%; Score 1200; DB 5; Length 225;  
Best Local Similarity 98.7%; Pred. No. 3.9e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD 60  
|||||  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMD 60

QY 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 120  
|||||  
Db 61 LSPQFGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGVTVTQDSSLQDGCFTY 120

QY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
Db 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 60  
AAE28926  
ID AAE28926 standard; protein; 225 AA.  
XX AC AAE28926;  
XX DT 27-DEC-2002 (first entry)  
XX DE Diacosoma sp. drFP583 (NFP-6) mutant protein, (V71M+V105A+Y120H+S197T).  
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutant;  
KW mutein.  
XX Diacosoma sp.  
OS Synthetic.  
OS  
XX Key Location/Qualifiers  
FH Misc-difference 71 /note= "Wild-type Val substituted with Met"  
FT Misc-difference 105 /note= "Wild-type Val substituted with Ala"  
FT Misc-difference 120 /note= "Wild-type Tyr substituted with His"  
FT Misc-difference 197 /note= "Wild-type Ser substituted with Thr"  
FT  
XX WO200268459-A2.  
XX 06-SEP-2002.  
XX 20-FEB-2002; 2002WO-US005749.



21-FEB-2001; 2001US-0270983P.  
04-DEC-2001; 2001US-00006922.  
(CLON-) CLONTECH LAB INC.  
Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
WPI; 2002-691654/74.  
New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
analyte detection assays or fluorescence activated cell sorting  
applications.  
Disclosure; Page; 80pp; English.  
The invention relates to nucleic acid molecules encoding non-aggregating  
chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
useful in analyte detection assays, as colouring agents, as markers in  
recombinant DNA applications, as sunscreens or filters, in fluorescence  
resonance energy transfer (FRET) applications, as biosensors in  
prokaryotic and eukaryotic cells, in screening assays, as second  
messenger detectors, in fluorescence activated cell sorting applications,  
in protease cleavage assays or as fluorescent timers. The present  
sequence is Discosoma sp. drfp583 (NFP-6) mutant protein of the  
invention. Note: This sequence is not shown in the specification, but is  
derived from Discosoma sp. drfp583 (NFP-6) wild-type protein shown as SEQ  
ID NO:8 (AAE28833) in page 70-71 of the specification  
Sequence 225 AA;

Query Match	98.8%	Score 1199;	DB 5;	Length 225;
Best Local Similarity	98.2%	Pred. No. 5e-126;		
Matches 221; Conservative	3;	Mismatches 1;	Indels 0;	Gaps 0;

[illegible]

RESULT	61
ADH34505	
ID	ADH34505 standard; protein; 225 AA.
XX	
AC	ADH34505;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	Discosoma sp. DsRed mutant S148C/I165N/S203A.
XX	
KW	Chromoprotein; fluorescent protein; CP; FP; Interconverted mutant;
KW	Cnidarian; Anthozoan; labelling; colouring agents; pigment;
KW	analyte detection assay; selectable marker; screen; selective filter;
KW	fluorescence resonance energy transfer; FRET; biosensor;
KW	whole cell marker; second messenger detector; in vivo marker;
KW	fluorescence activated cell sorting; fluorescent timer;
KW	red fluorescent protein; DsRed; mutant; mutain.
XX	
OS	Synthetic.
OS	Discosoma sp.

XX	Key	Location/Qualifiers
FH	Misc-difference	146
FT		/note= "Cys replaces wild-type Ser. Corresponds to GFP
FT		residue 148"
FT	Misc-difference	161
FT		/note= "Asn replaces wild-type Ile. Corresponds to GFP
FT		residue 165"
FT	Misc-difference	197
FT		/note= "Ala replaces wild-type Ser. Corresponds to GFP
FT		residue 203"
XX		
PN	WO2003057833-A2.	
XX		
PD	17-JUL-2003.	
XX		
PF	23-DEC-2002; 2002WO-US041418.	
XX		
PR	26-DEC-2001; 2001US-0343128P.	
XX		
PA	(CLON-) CLONTECH LAB INC.	
XX		
PI	Bulina ME, Chudakov D, Lukyanov KA;	
XX		
DR	WPI; 2003-607998/57.	
XX		
PT	Novel nucleic acid encoding interconverted mutant of chromo-or	
PT	fluorescent protein which are useful as biosensors, coloring agents.	
XX		
SS	Example 1; Page; 56pp; English.	

The invention relates to interconverted mutants of chromoproteins (CP) or fluorescent proteins (FP) and nucleic acids encoding them. The mutant is derived from a Cnidarian species, preferably a non-bioluminescent Cnidarian species, and most preferably an Anthozoan species. The invention is based on the finding that although green fluorescent protein (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of homology, there are certain positions (referred to as 148, 165, 167 and 203; numbering corresponds to GFP) that are occupied by noticeably different residues in the two types of proteins. Mutagenesis of the residues in these key positions in, for example, a fluorescent protein, to those found in a chromoprotein is therefore proposed to confer chromoprotein activity on the fluorescent protein mutant, with chromoproteins being able to be converted into fluorescent proteins in a similar manner. The invention also relates to expression constructs, vectors, host cells and host cell progeny comprising a nucleic acid of the invention; the recombinant production of an interconverted chromoprotein or fluorescent protein mutant; and antibodies specific for interconverted mutant proteins of the invention. The interconverted mutants are useful in any application that employs a chromoprotein or fluorescent protein. Fluorescent protein mutants having chromoprotein activity can be useful as colouring agents in, for example, food compositions, pharmaceuticals, cosmetics and living organisms. Proteins with chromoprotein activity are also useful as labels in biological analyte detection assays, as selectable markers in recombinant DNA applications (e.g. the production of transgenic cells and organisms), and are also useful as screens and selective filters. Chromoprotein mutants having fluorescent protein activity useful in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, as markers of whole cells to detect changes in multicellular reorganisation and migration, as second messenger detectors, as *in vivo* markers in animals (e.g., transgenic animals), in fluorescence activated cell sorting applications, in protease cleavage assays, and in assays to determine the phospholipid composition in biological membranes. Proteins with fluorescent protein activity can also be used as fluorescent timers, where the switch of one fluorescent colour to another (e.g., green to red) is concomitant with the ageing of the protein and is useful for determination of the activation or deactivation of gene expression. The present sequence represents a *Discosoma* sp. red fluorescent protein DsRed mutant generated in an example of the invention. The present sequence is not shown in the specification, but was derived from the wild-type DsRed sequence (ADH34489) shown in Fig 1 and the information provided on page 42.



XX Sequence 225 AA;  
SQ  
Query Match 98.8%; Score 1199; DB 7; Length 225;  
Best Local Similarity 98.7%; Pred. No. 5e-126;  
Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 1 MRSSKNVTKEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI 60  
DB 1 MRSSKNVTKEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI 60  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYVVD SKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVD AKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 62  
ABP56684  
ABP56684 standard; protein; 240 AA.  
XX AC ABP56684;  
XX DT 25-MAR-2003 (first entry)  
XX DE Discosoma red fluorescent protein variant F125L.  
XX KM Mammalian codon optimised Discosoma red fluorescent protein; variant;  
XX KM Discosoma; red fluorescent protein; directed evolution; mutant.  
XX OS Discosoma sp.  
XX OS Synthetic.  
XX PN WO200294992-A2.  
XX PD 28-NOV-2002.  
XX PF 20-MAY-2002; 2002WO-US015968.  
XX PR 18-MAY-2001; 2001US-0291871P.  
XX PA (RIGE-) RIGEL PHARM INC.  
XX PI Peelle B;  
XX DR WPI; 2003-120798/11.  
XX PT New Discosoma red fluorescent protein, useful for functional screens as a  
XX PT reporter for gene transcription, for target characterization and  
XX PT localization of fusion proteins, or for scaffolds for protein and peptide  
XX PT libraries.  
XX PS Claim 15; Page; 22pp; English.  
XX CC The present invention describes an isolated Discosoma red fluorescent  
XX CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
XX CC S1), with one or more point mutations at amino acid position N24, F125,  
XX CC K164, or M183. Also described: (1) a fusion protein comprising (I); (2)  
XX CC an isolated nucleic acid encoding (I); (3) a vector comprising the  
XX CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
XX CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
XX CC encoding (I); and (6) methods of making a fluorescent variant. The  
XX CC Discosoma red fluorescent proteins are useful for functional screens as a  
XX CC reporter for gene transcription (e.g. as a fusion protein), for target  
XX CC characterisation and localisation of fusion proteins, or for scaffolds  
XX CC for protein and peptide libraries. The fluorescent proteins can also be

CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
XX claims  
XX Sequence 240 AA;  
SQ  
Query Match 98.8%; Score 1199; DB 6; Length 240;  
Best Local Similarity 98.7%; Pred. No. 5.5e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 MRSSKNVTKEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI 60  
DB 2 VRSSKNVTKEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVTKLVTKGGLPFAWDI 61  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 62 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 121  
OY 121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 122 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
OY 181 YMAKKPVQLPGYYVVD SKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 182 YMAKKPVQLPGYYVVD SKLDITSHNEDYTIIVEQYERTEGRHHLFL 226  
RESULT 63  
ABP56683  
ABP56683 standard; protein; 240 AA.  
XX AC ABP56683;  
XX DT 25-MAR-2003 (first entry)  
XX DE Discosoma red fluorescent protein variant K164M.  
XX KM Mammalian codon optimised Discosoma red fluorescent protein; variant;  
XX KM Discosoma; red fluorescent protein; directed evolution; mutant.  
XX OS Discosoma sp.  
XX OS Synthetic.  
XX PN WO200294992-A2.  
XX PD 28-NOV-2002.  
XX PF 20-MAY-2002; 2002WO-US015968.  
XX PR 18-MAY-2001; 2001US-0291871P.  
XX PA (RIGE-) RIGEL PHARM INC.  
XX PI Peelle B;  
XX DR WPI; 2003-120798/11.  
XX PT New Discosoma red fluorescent protein, useful for functional screens as a  
XX PT reporter for gene transcription, for target characterization and  
XX PT localization of fusion proteins, or for scaffolds for protein and peptide  
XX PT libraries.  
XX PS Claim 14; Page; 22pp; English.  
XX CC The present invention describes an isolated Discosoma red fluorescent



CC protein (1) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, P125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims  
CC  
SQ Sequence 240 AA;

Query Match 98.8%; Score 1199; DB 6; Length 240;  
Best Local Similarity 98.7%; Pred. No. 5.5e-126;  
Matches 222; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db :|||||  
2 VRSSKNVYKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 61  
QY 61 LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 120  
Db |||||  
62 LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFTY 121  
QY 121 KVKFAGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHHYLVFKSI 180  
Db |||||  
122 KVKFAGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHHYLVFKSI 181  
QY 181 YMAKKEVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db |||||  
182 YMAKKEVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 64  
ADH34504  
ID ADH34504 standard; protein; 225 AA.  
XX  
AC ADH34504;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Discosoma sp. DsRed mutant S148A/I165S/K167M/S203A.  
XX  
XX Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRET; biosensor;  
KW whole cell marker; second messenger detector; in vivo marker;  
KW fluorescence activated cell sorting; fluorescent timer;  
KW red fluorescent protein; DsRed; mutant; mutein.  
XX  
XX Synthetic.  
OS Discosoma sp.  
OS  
XX  
FH Key location/Qualifiers  
FT Misc-difference 146  
FT /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 148"  
FT Misc-difference 161  
FT /note= "Ser replaces wild-type Ile. Corresponds to GFP

FT

FT residue 165"  
FT Misc-difference 163  
FT /note= "Met replaces wild-type Lys. Corresponds to GFP  
FT residue 167"  
FT Misc-difference 197  
FT /note= "Ala replaces wild-type Ser. Corresponds to GFP  
FT residue 203"  
XX  
PN WO2003057833-A2.  
XX  
PD 17-JUL-2003.  
XX  
XX 23-DEC-2002; 2002WO-US041418.  
PF 26-DEC-2001; 2001US-0343128P.  
PR  
XX  
XX (CLON-) CLONTECH LAB INC.  
PA  
PI Bulina ME, Chudakov D, Lukyanov KA;  
XX  
XX WPI; 2003-607998/57.  
DR  
XX

Novel nucleic acid encoding interconverted mutant of chromo-or  
fluorescent protein which are useful as biosensors, coloring agents.  
Example 1; Page; 56pp; English.

The invention relates to interconverted mutants of chromoproteins (CP) or  
fluorescent proteins (FP) and nucleic acids encoding them. The mutant is  
derived from a Cnidarian species, preferably a non-bioluminescent  
Cnidarian species, and most preferably an Anthozoan species. The  
invention is based on the finding that although green fluorescent protein  
(GFP)-like chromoproteins and fluorescent proteins exhibit some degree of  
homology, there are certain positions (referred to as 148, 165, 167 and  
203; numbering corresponds to GFP) that are occupied by noticeably  
different residues in the two types of proteins. Mutagenesis of the  
residues in these key positions in, for example, a fluorescent protein,  
to those found in a chromoprotein is therefore proposed to confer  
chromoprotein activity on the fluorescent protein mutant, with  
chromoproteins being able to be converted into fluorescent proteins in a  
similar manner. The invention also relates to expression constructs,  
vectors, host cells and host cell progeny comprising a nucleic acid of  
the invention; the recombinant production of an interconverted  
chromoprotein or fluorescent protein mutant; and antibodies specific for  
interconverted mutant proteins of the invention. The interconverted  
mutants are useful in any application that employs a chromoprotein or  
fluorescent protein. Fluorescent protein mutants having chromoprotein  
activity can be useful as colouring agents in, for example, food  
compositions, pharmaceuticals, cosmetics and living organisms. Proteins  
with chromoprotein activity are also useful as labels in biological  
analyte detection assays, as selectable markers in recombinant DNA  
applications (e.g. the production of transgenic cells and organisms), and  
are also useful as sunscreens and selective filters. Chromoprotein  
mutants having fluorescent protein activity useful in fluorescence  
resonance energy transfer (FRET) applications, as biosensors in  
prokaryotic and eukaryotic cells, as markers of whole cells to detect  
changes in multicellular reorganisation and migration, as second  
messenger detectors, as in vivo markers in animals (e.g., transgenic  
animals), in fluorescence activated cell sorting applications, in  
protease cleavage assays, and in assays to determine the phospholipid  
composition in biological membranes. Proteins with fluorescent protein  
activity can also be used as fluorescent timers, where the switch of one  
fluorescent colour to another (e.g., green to red) is concomitant with  
the ageing of the protein and is useful for determination of the  
activation or deactivation of gene expression. The present sequence  
represents a Discosoma sp. red fluorescent protein DsRed mutant generated  
in an example of the invention. The present sequence is not shown in the  
specification, but was derived from the wild-type DsRed sequence  
(ADH34489) shown in Fig 1 and the information provided on page 42.

SQ Sequence 225 AA;

Query Match 98.5%; Score 1196; DB 7; Length 225;



Best local similarity 98.2%; Pred. No. 1.1e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MRSSKNVTKEMRPFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKVTKGGLPFAWDI 60  
Db 1 MRSSKNVTKEMRPFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKVTKGGLPFAWDI 60  
Qy 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120  
Qy 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGESHMALTKDKGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTVIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTVIVEQYERTEGRHHLFL 225

RESULT 65  
ADO78068  
ID ADO78068 standard; protein; 236 AA.  
XX ADO78068;  
AC  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Corallimorpharia red fluorescent protein mutant RedI.  
XX  
KW fluorescent protein; red fluorescent protein; green fluorescent protein;  
KW gene expression detection; regeneration; multiple labeling system;  
KW in vivo marker; microinjection assay; taxonomic marker; colour indicator;  
KW food additive; cosmetic; Actinodiscus; Discosoma; muten; mutant; RedI;  
KW coral-like anemone.

XX Corallimorpharia.  
OS  
XX US2004110225-A1.  
PN  
XX 10-JUN-2004.  
PD  
XX  
PF 09-DEC-2002; 2002US-00314936.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
XX  
PI Gibbs PDL, Carter RW, Schmale MC;  
XX  
XX WPI; 2004-467659/44.  
DR N-PSDB; ADO78067.  
DR  
XX  
XX  
PT New first mutant fluorescent protein having an optical property  
PT relatively different to a corresponding optical property produced by a  
PT wild-type red or green fluorescent protein, useful for detecting  
PT expression of a gene.  
XX  
XX  
PS Claim 6; SEQ ID NO 2; 30pp; English.  
XX  
CC The invention describes a first mutant fluorescent protein (I) comprising  
CC an amino acid sequence that differs from that of a corresponding wild-  
CC type red fluorescent protein by an amino acid substitution, where the  
CC first mutation fluorescent protein mutant has an optical property that  
CC differs relative to the corresponding optical property produced by the  
CC wild-type red or green fluorescent protein. (I) is useful for detecting  
CC expression of gene. The polynucleotide (II) encoding (I) is useful for  
CC detecting expression of a gene which involves introducing (II) into a  
CC cell or organism, allowing the cell to replicate, and detecting  
CC expression of the nucleic acid by emission of fluorescent light. The  
CC expression of the nucleic acid expression is detected in vivo or in

CC vitro. (II) is useful for regenerating an animal such as a zebrafish  
CC which contains (II). (I) is useful as a marker for detecting expression  
CC of a gene, in biochemical assays, and as reagents. (I) is also useful in  
CC multiple labeling systems, as in vivo markers such as in mRNA  
CC microinjection assays, and as taxonomic markers for studies of genetics,  
CC colour indicators in diagnostic kits, coloured food additives, and  
CC cosmetics ingredients. (I) is useful in research for up or down  
CC regulation, to monitor promoter activity, to allow longer term monitoring  
CC and to localise proteins. (I) has enhanced properties such as  
CC substantially enhanced fluorescence and reduced toxicity. This is the  
CC amino acid sequence of RedI, a mutant of the red fluorescent protein  
CC isolated from an aquatic species believed to be either an Actinodiscus or  
CC Discosoma species.

Seq Sequence 236 AA;  
Query Match 98.5%; Score 1196; DB 8; Length 236;  
Best local similarity 98.2%; Pred. No. 1.2e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRSSKNVTKEMRPFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKVTKGGLPFAWDI 60  
Db 1 MRSSKNVTKEMRPFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKVTKGGLPFAWDI 60  
Qy 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDEGVVTVTQDSSLQDGCFTY 120  
Qy 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDKDGGHYLVEFKSI 180  
Qy 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTVIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTVIVEQYERTEGRHHLFL 225

RESULT 66  
ADO78070  
ID ADO78070 standard; protein; 236 AA.  
XX  
XX ADO78070;  
AC  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Corallimorpharia red fluorescent protein mutant RedI.  
XX  
KW fluorescent protein; red fluorescent protein; green fluorescent protein;  
KW gene expression detection; regeneration; multiple labeling system;  
KW in vivo marker; microinjection assay; taxonomic marker; colour indicator;  
KW food additive; cosmetic; Actinodiscus; Discosoma; muten; mutant; RedI;  
KW coral-like anemone.  
XX  
XX Corallimorpharia.  
OS  
XX US2004110225-A1.  
PN  
XX 10-JUN-2004.  
PD  
XX  
PF 09-DEC-2002; 2002US-00314936.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
XX  
PI Gibbs PDL, Carter RW, Schmale MC;  
XX  
XX WPI; 2004-467659/44.  
DR N-PSDB; ADO78069.  
DR  
XX  
PT New first mutant fluorescent protein having an optical property



PT relatively different to a corresponding optical property produced by a  
PT wild-type red or green fluorescent protein, useful for detecting  
PT expression of a gene.  
PS Claim 7; SEQ ID NO 4; 30pp; English.  
XX  
CC The invention describes a first mutant fluorescent protein (I) comprising  
CC an amino acid sequence that differs from that of a corresponding wild-  
CC type red fluorescent protein by an amino acid substitution, where the  
CC first mutation fluorescent protein mutant has an optical property that  
CC differs relative to the corresponding optical property produced by the  
CC wild-type red or green fluorescent protein. (I) is useful for detecting  
CC expression of gene. The polynucleotide (II) encoding (I) is useful for  
CC detecting expression of a gene which involves introducing (II) into a  
CC cell or organism, allowing the cell to replicate, and detecting  
CC expression of the nucleic acid by emission of fluorescent light. The  
CC expression of the nucleic acid expression is detected in vivo or in  
CC vitro. (II) is useful for regenerating an animal such as a zebrafish  
CC which contains (II). (I) is useful as a marker for detecting expression  
CC of a gene, in biochemical assays, and as reagents. (I) is also useful in  
CC multiple labeling systems, as in vivo markers such as in mRNA  
CC microinjection assays, and as taxonomic markers for studies of genetics,  
CC colour indicators in diagnostic kits, coloured food additives, and  
CC cosmetics ingredients. (I) is useful in research for up or down  
CC regulation, to monitor promoter activity, to allow longer term monitoring  
CC and to localise proteins. (I) has enhanced properties such as  
CC substantially enhanced fluorescence and reduced toxicity. This is the  
CC amino acid sequence of RedII, a mutant of the red fluorescent protein  
CC isolated from an aquatic species believed to be either an Actinodiscus or  
CC Discosoma species.  
XX  
SQ Sequence 236 AA;  
Query Match 98.5%; Score 1196; DB 8; Length 236;  
Best Local Similarity 98.2%; Pred. No. 1.2e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKYTKGGLPFAWDI 60  
DB 1 MSCSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKYTKGGLPFAWDI 60  
QY 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGCVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKTI 180  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNKDYTIIVEQYERTEGRHHLFL 225  
RESULT 67  
ADZ84223  
ID ADZ84223 standard; protein; 236 AA.  
XX  
AC ADZ84223;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE Actinodiscus/Discosoma RFP mutant Red I, SEQ ID NO:2.  
XX  
KW Red fluorescent protein; Red I; mutein.  
XX  
OS Corallimorpharia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 232 /note= "This residue is Asn in the mutant Ac/DsRFP Red  
FT II"

XX  
PN US2005100954-A1.  
XX  
PD 12-MAY-2005.  
XX  
PF 23-DEC-2004; 2004US-00021014.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
XX  
PI Gibbs PDL, Carter RW, Schmale MC;  
XX  
DR MPI; 2005-345399/35.  
DR N-PSDB; ADZ84222.  
XX  
PT Novel mutant fluorescent protein comprising amino acid sequence that  
PT differs from wild-type red or green fluorescent protein by amino acid  
PT substitutions, useful as markers for detecting desired gene expression.  
PS Claim 6; SEQ ID NO 2; 29pp; English.  
XX  
CC The invention relates to mutant red and green fluorescent proteins which  
CC have higher fluorescent intensities compared to the wild-type proteins  
CC from which they are derived. The invention also discloses nucleic acids  
CC encoding the mutant fluorescent proteins, and transgenic animals  
CC comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
CC of the invention, Red I (ADZ84223) and Red II (ADZ84225), are derived  
CC from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
CC mushroom coral believed to be either an Actinodiscus or Discosoma  
CC species, while the mutant green fluorescent proteins (GFP), Green I  
CC (ADZ84227) and Green (ADZ84229), are derived from a novel wild-type GFP  
CC (known as McGFP) isolated from the great star coral Montastraea  
CC cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the  
CC wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce  
CC random mutations, cloning the PCR products into bacterial expression  
CC vectors, and selecting the transformed bacteria having higher relative  
CC fluorescence compared to those expressing the wild-type proteins. The  
CC mutant fluorescent proteins of the invention may be used in a wide range  
CC of applications. They are useful as markers for detecting the expression  
CC of a gene of interest, or can be used as in vivo markers in mRNA  
CC microinjection assays in transgenic animals. They can be used in  
CC expression systems (e.g., as components of fusion proteins), in multiple  
CC labeling systems, or can be used in biochemical assays and as reagents.  
CC They may additionally be used as taxonomic markers for studies of  
CC cnidarian genetics, as color indicators in diagnostic kits, as colored  
CC food additives and as cosmetic ingredients. The present sequence  
CC represents the mutant Actinodiscus/Discosoma red fluorescent protein Red  
CC I.  
XX  
SQ Sequence 236 AA;  
Query Match 98.5%; Score 1196; DB 9; Length 236;  
Best Local Similarity 98.2%; Pred. No. 1.2e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKYTKGGLPFAWDI 60  
DB 1 MSCSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKYTKGGLPFAWDI 60  
QY 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFQYGSKYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGCVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
DB 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKTI 180  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVYVDSKLDITSHNKDYTIIVEQYERTEGRHHLFL 225



RESULT 68

AD284225

ID

AD284225

standard; protein; 236 AA.

XX

AC

AD284225;

XX

DT

14-JUL-2005

(first entry)

XX

DE

Actinodiscus/Discosoma RFP mutant Red II, SEQ ID NO:4.

XX

KW

Red fluorescent protein; Red II; mutein.

XX

OS

CoralImorpharia.

OS

Synthetic.

FH

Key

Location/Qualifiers

PT

Misc-difference

232

/note= "This residue is Asp in the mutant Ac/DsRFP Red I"

XX

PN

US2005100954-A1.

XX

PD

12-MAY-2005.

XX

PF

23-DEC-2004; 2004US-00021014.

XX

PR

09-DEC-2002; 2002US-00314936.

XX

PA

(GIBB/) GIBBS P D L.

PA

(CART/) CARTER R W.

PA

(SCHM/) SCHMALE M C.

XX

PI

Gibbs PDL, Carter RW, Schmale MC;

XX

DR

WPI; 2005-345399/35.

XX

DR

N-PSDB; AD284224.

XX

PT

Novel mutant fluorescent protein comprising amino acid sequence that differs from wild-type red or green fluorescent protein by amino acid substitutions, useful as markers for detecting desired gene expression.

XX

PS

Claim 7; SEQ ID NO 4; 29pp; English.

XX

CC

The invention relates to mutant red and green fluorescent proteins which have higher fluorescent intensities compared to the wild-type proteins from which they are derived. The invention also discloses nucleic acids encoding the mutant fluorescent proteins, and transgenic animals comprising such a nucleic acid. The mutant red fluorescent proteins (RFP) of the invention, Red I (AD284223) and Red II (AD284225), are derived from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a mushroom coral believed to be either an Actinodiscus or Discosoma species, while the mutant green fluorescent proteins (GFP), Green I (AD284227) and Green (AD284229), are derived from a novel wild-type GFP (known as McGFP) isolated from the great star coral Montastraea cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce random mutations, cloning the PCR products into bacterial expression vectors, and selecting the transformed bacteria having higher relative fluorescence compared to those expressing the wild-type proteins. The mutant fluorescent proteins of the invention may be used in a wide range of applications. They are useful as markers for detecting the expression of a gene of interest, or can be used as in vivo markers in mRNA microinjection assays in transgenic animals. They can be used in labeling systems (e.g., as components of fusion proteins), in multiple expression systems, or can be used in biochemical assays and as reagents. They may additionally be used as taxonomic markers for studies of cnidarian genetics, as color indicators in diagnostic kits, as colored food additives and as cosmetic ingredients. The present sequence represents the mutant Actinodiscus/Discosoma red fluorescent protein Red II. This protein was found to have a fluorescent intensity at least 50% greater than the mutant RFP Red I (AD284223).

SEQ

Sequence

236 AA;

Query Match

98.5%;

Score 1196;

DB 9;

Length 236;

Best Local Similarity

98.2%;

Pred. No. 1.2e-125;

Matches 221;

Conservative 2;

Mismatches 2;

Indels 0;

Gaps 0;

QY

1

MRSSKNVTKEMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI

60

DB

1

MSCSKNVTKEMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI

60

QY

61

LSPQFYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY

120

DB

61

LSPQFYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGVVTVTQDSSLQDGCFTY

120

QY

121

KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEPKSI

180

DB

121

KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKTI

180

QY

181

YMAKKPVQLPGYYVVDSDKLDTSHNEDYTVIEQYERTEGRHHHLFL

225

DB

181

YMAKKPVQLPGYYVVDSDKLDTSHNKDYTVIEQYERTEGRHHHLFL

225

RESULT 69

ADQ59562

ID

ADQ59562

standard; protein; 284 AA.

XX

AC

ADQ59562;

DT

07-OCT-2004

(first entry)

XX

DE

Discosoma fluorescent protein drFP583.

XX

KW

optically detectable signal; protein fragment complementation assay; PCA; molecular interaction detection; biomolecular interaction detection; fluorescent protein; multi-colour PCA; drug discovery; target validation; high-throughput screening; high-content screening; pathway mapping; drug mechanism-of-action study; biosensor; diagnostic;

XX

KW

fluorescent protein; spectral shift.

XX

OS

Discosoma.

XX

PN

US2004137528-A1.

XX

PD

15-JUL-2004.

XX

PF

01-DEC-2003; 2003US-00724178.

XX

PR

02-FEB-1998; 98US-00017412.

PR

07-FEB-2000; 2000US-00499464.

PR

24-MAY-2002; 2002US-00154758.

PR

29-JAN-2003; 2003US-00353090.

PR

09-APR-2003; 2003US-0461133P.

XX

PA

(MICH/) WATSON MICHNICK S W.

PA

(MACD/) MACDONALD M L.

PA

(LAME/) LAMERDIN J.

XX

PI

Watson Michnick SW, Macdonald ML, Lamerdin J;

XX

DR

WPI; 2004-533363/51.

XX

PT

Composition useful in protein fragment complementation assays for drug discovery and high-throughput screening, comprising complementary fragments of protein or mutant protein, generating optically detectable signal when associated.

XX

PS

Disclosure; SEQ ID NO 12; 34pp; English.

XX

CC

The invention describes a composition (I) comprising complementary fragments of a protein or mutant protein, where the fragments generate an optically detectable signal when associated, and each of the mutant protein fragments is fused to a separate molecule. Also described are:



CC protein fragment complementation assays (PCAs) for detection of molecular  
CC interactions, involving reassembling separate fragments from an optically  
CC detectable protein, and detecting the reassembly by units of  
CC reconstruction of the fragments is operated by the interaction of molecular  
CC domains fused to each fragment, and is independent of other molecular  
CC processes; detecting biomolecular interaction, involving selecting an  
CC appropriate optically detectable protein, effecting fragmentation of the  
CC optically detectable protein such that the fragmentation results in  
CC reversible loss of protein function, fusing or attaching fragments of the  
CC optically detectable protein separately to other molecules, reassociating  
CC the protein fragments through interactions of the molecules that are  
CC fused or attached to the fragments, and detecting the resulting optical  
CC signal; designing and engineering of PCAs based on fluorescent protein;  
CC and a method and composition for the construction of multi-color PCAs.  
CC (I) is useful in PCAs and other assays for drug discovery, target  
CC validation, high-throughput screening, high-content screening, pathway  
CC mapping, drug mechanism-of-action studies, biosensors and diagnostics.  
CC (I) is useful for engineering different colour PCAs for a variety of  
CC applications in biology and biotechnology. This is the amino acid  
CC sequence of a Discosoma fluorescent protein used in a comparison of  
CC fluorescent proteins with Aequorea green fluorescent protein, variants of  
CC which can be created with altered excitation and emission wavelength  
CC relative to wild type GFP.

XX SQ Sequence 284 AA;

Query Match 98.5%; Score 1195.5; DB 8; Length 284;  
Best Local Similarity 99.1%; Pred. No. 1.8e-125;  
Matches 224; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
|||  
DB 59 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 118  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
|||  
DB 119 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 178  
OY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
DB 179 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 238  
OY 181 YMAKK-PVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
DB 239 YMAKKAPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 284

RESULT 70

ABP56681  
ID ABP56681 standard; protein; 240 AA.

XX AC ABP56681;

XX DT 25-MAR-2003 (first entry)

XX DE Discosoma red fluorescent protein variant F125L and N24S.

KW Mammalian codon optimised Discosoma red fluorescent protein; variant;  
KW Discosoma; red fluorescent protein; directed evolution; mutant.

XX OS Discosoma sp.  
XX OS Synthetic.

XX PN WO200294992-A2.

XX PD 28-NOV-2002.

XX PF 20-MAY-2002; 2002WO-US015968.

XX PR 18-MAY-2001; 2001US-0291871P.

XX PA (RIGE-) RIGEL PHARM INC.

XX  
PI Peelle B;  
XX  
DR WPI; 2003-120798/11.  
XX  
PT New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.  
XX  
PS Claim 11; Page; 22pp; English.

CC The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (I); (2)  
CC an isolated nucleic acid encoding (I); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (I); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims

XX SQ Sequence 240 AA;

Query Match 98.4%; Score 1194; DB 6; Length 240;  
Best Local Similarity 98.2%; Pred. No. 2e-125;  
Matches 221; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVRMEGTVSGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAWDI 61  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
|||  
DB 62 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 121  
OY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||  
DB 122 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
OY 181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||  
DB 182 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 71

ADH34506  
ID ADH34506 standard; protein; 225 AA.

XX AC ADH34506;

XX DT 11-MAR-2004 (first entry)

XX DE Discosoma sp. DsRed mutant DsRed-NF S148C/I165N/K167M/S203A.

KW Chromoprotein; fluorescent protein; CP; FP; interconverted mutant;  
KW Cnidarian; Anthozoan; labelling; colouring agents; pigment;  
KW analyte detection assay; selectable marker; sunscreen; selective filter;  
KW fluorescence resonance energy transfer; FRRET; biosensor;



KW		whole cell marker; second messenger detector; in vivo marker;
KV		fluorescence activated cell sorting; fluorescent timer;
KW		red fluorescent protein; Dered; non-fluorescent mutant; Dered-NF; mutant;
KX		muteln.
XX		
OS	Synthetic.	
OS	Dicosoma sp.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	146
PT	/note= "Cys replaces wild-type Ser. Corresponds to GFP	
FT	residue 148"	
FT	Misc-difference	161
FT	/note= "Asn replaces wild-type Ile. Corresponds to GFP	
FT	residue 165"	
FT	Misc-difference	163
FT	/note= "Met replaces wild-type Lys. Corresponds to GFP	
FT	residue 167"	
FT	/note= "Ala replaces wild-type Ser. Corresponds to GFP	
FT	residue 203"	
XX		
PN	WO2003057833-A2.	
XX		
PD	17-JUL-2003.	
XX		
PB	23-DEC-2002; 2002WO-US041418.	
PR	26-DEC-2001; 2001US-0343128P.	
XX		
PA	(CLON-) CLONTECH LAB INC.	
XX		
PI	Bulina ME, Chudakov D, Lukyanov KA;	
XX		
DR	WPI; 2003-607998/57.	

Novel nucleic acid encoding interconverted mutant of chromo-or  
fluorescent protein which are useful as biosensors, coloring agents.  
Example 1; Page; 56pp; English.

The invention relates to interconverted mutants of chromoproteins (CP) or fluorescent proteins (FP) and nucleic acids encoding them. The mutant is derived from a Cnidarian species, preferably a non-bioluminescent Cnidarian species, and most preferably an Anthozoan species. The invention is based on the finding that although green fluorescent protein (GFP)-like chromoproteins and fluorescent proteins exhibit some degree of homology, there are certain positions (referred to as 148, 165, 167 and 203; numbering corresponds to GFP) that are occupied by noticeably different residues in the two types of proteins. Mutagenesis of the residues in these key positions in, for example, a fluorescent protein, to those found in a chromoprotein is therefore proposed to confer chromoprotein activity on the fluorescent protein mutant, with chromoproteins being able to be converted into fluorescent proteins in a similar manner. The invention also relates to expression constructs, vectors, host cells and host cell progeny comprising a nucleic acid of the invention; the recombinant production of an interconverted chromoprotein or fluorescent protein mutant; and antibodies specific for interconverted mutant proteins of the invention. The interconverted mutants are useful in any application that employs a chromoprotein or fluorescent protein. Fluorescent protein mutants having chromoprotein activity can be useful as colouring agents in, for example, food compositions, pharmaceuticals, cosmetics and living organisms. Proteins with chromoprotein activity are also useful as labels in biological analyte detection assays, as selectable markers in recombinant DNA applications (e.g. the production of transgenic cells and organisms), and are also useful as sunscreens and selective filters. Chromoprotein mutants having fluorescent protein activity useful in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, as markers of whole cells to detect changes in multicellular reorganisation and migration, as second messenger detectors, as in vivo markers in animals (e.g., transgenic animals), in fluorescence activated cell sorting applications, in

CC protease cleavage assays, and in assays to determine the phospholipid  
CC composition in biological membranes. Proteins with fluorescent protein  
CC activity can also be used as fluorescent timers, where the switch of one  
CC fluorescent colour to another (e.g., green to red) is concomitant with  
CC the ageing of the protein and is useful for determination of the  
CC activation or deactivation of gene expression. The present sequence  
CC represents a non-fluorescent *Drosophila* sp. red fluorescent protein DsRed  
CC mutant, DsRed-NF, which has chromoprotein activity and which was  
CC generated in an example of the invention. The present sequence is not  
CC shown in the specification, but was derived from the wild-type DsRed  
CC sequence (ADH34489) shown in Fig 1 and the information provided on page  
CC 42.  
SQ Sequence 225 AA;  
XX

Query Match	98.3%;	Score 1193;	DB 7;	Length 225;
Best Local Similarity	98.2%;	Pred. No. 2.4e-125;		
Matches 221; Conservative	1;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1	MRSSKNV	IKENR	FKVR	MEGT	VNGH	FEI	EGEG	EGGR	PYEG	HTV	KL	KVTK	GG	BL	PF	AW	DI	60
DB	1	MRSSKNV <td>IKENR <td>FKVR <td>MEGT <td>VNGH <td>FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	IKENR <td>FKVR <td>MEGT <td>VNGH <td>FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	FKVR <td>MEGT <td>VNGH <td>FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	MEGT <td>VNGH <td>FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td></td></td>	VNGH <td>FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td></td>	FEI <td>EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td></td>	EGEG <td>EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td></td>	EGGR <td>PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td></td>	PYEG <td>HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td></td>	HTV <td>KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td></td>	KL <td>KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td></td>	KVTK <td>GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td></td>	GG <td>BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td></td>	BL <td>PF <td>AW <td>DI</td> <td>60</td> </td></td>	PF <td>AW <td>DI</td> <td>60</td> </td>	AW <td>DI</td> <td>60</td>	DI	60

QY	61	LS	PQ	F	O	G	S	K	V	Y	K	H	P	A	D	I	P	Y	K	L	S	F	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	T	T	O	D	S	L	O	D	G	C	F	I	Y	120	
Db	61	LS	P	O	F	O	G	S	K	V	Y	K	H	P	A	D	I	P	Y	K	L	S	F	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	T	T	O	D	S	L	O	D	G	C	F	I	Y	120

QY	121 KYKFIGVNPSPDGPVMQKTGMWEASTERLYPRDGLVKG EIHAKLKDGGHYLVEEKSI 180
Db	121 KYKFIGVNPSPDGPVMQKTGMWEACTERLYPRDGLVKG ENHMAKLKDGGHYLVEEKSI 180

```
QY      181 YMAKKPVLQEGYYVDSKLDITSHNEDYTIVEQYERTEGRHHFL 225
        ||||| | | | | |
        |||||:|||||
        |||||
Db      181 YMAKKPVLQEGYYVDADKLDITSHNEDYTIVEQYERTEGRHHFL 225
```

RESULT 72  
ABP56679  
ID ABP56679 standard; protein; 240 AA.

XX ABP56679;  
AC  
XX  
XX 25-MAR-2003 (first entry)  
DT  
XX  
DE Diaposoma red fluorescent protein variant F125L and M183K

AA Mammalian codon optimised Discosoma red fluorescent protein; variant;  
KW Discosoma; red fluorescent protein; directed evolution; mutant.

OS Discosoma sp.  
OS Synthetic.

PN WO200294992-A2.

PD 28-NOV-2002.

20-MAY-2002; 2002WO-US015968.

PR 18-MAY-2001; 2001US-0291871P.

PA (RIGE-) RIGEL PHARM INC.

PI Peelie B;

DR WPI; 2003-120798/11.

PT New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.

PS Claim 8; Page; 22pp; English.



CC The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims  
XX  
SQ Sequence 240 AA;

Query Match 98.3%; Score 1193; DB 6; Length 240;  
Best Local Similarity 98.2%; Pred. No. 2.6e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 61  
61 LSPQFGYSKVVYVADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
|||||  
DB 62 LSPQFGYSKVVYVADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCFTY 121  
121 KVKFIGNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
DB 122 KVKLIGNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
:|||||  
DB 182 YKAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 226

RESULT 73  
ABP56680  
ID ABP56680 standard; protein; 240 AA.  
AC ABP56680;  
XX  
DT 25-MAR-2003 (first entry)  
XX  
DE Discosoma red fluorescent protein variant F125V and M183K.  
XX  
KW Mammalian codon optimised Discosoma red fluorescent protein; variant;  
KW Discosoma; red fluorescent protein; directed evolution; mutant.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
PN WO200294992-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 20-MAY-2002; 2002WO-US015968.  
XX  
PR 18-MAY-2001; 2001US-0291871P.  
XX  
PA (RIGE-) RIGEL PHARM INC.  
XX

PI Peelle B;  
XX  
DR WPI; 2003-120798/11.  
XX  
PT New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.  
XX  
PS Claim 9; Page; 22pp; English.  
XX

CC The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (1); (2)  
CC an isolated nucleic acid encoding (1); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (1); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in the  
CC claims  
XX  
SQ Sequence 240 AA;

Query Match 98.2%; Score 1192; DB 6; Length 240;  
Best Local Similarity 98.2%; Pred. No. 3.4e-125;  
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
:|||||  
DB 2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 61  
61 LSPQFGYSKVVYVADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCFTY 120  
|||||  
DB 62 LSPQFGYSKVVYVADIPDYKKLSFPEGFKWERVMNFEDGGVTVTTQDSSLQDGCFTY 121  
121 KVKFIGNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
|||||  
DB 122 KVKVIGNPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181  
OY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225  
:|||||  
DB 182 YKAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 226

RESULT 74  
AAE28837  
ID AAE28837 standard; protein; 225 AA.  
AC AAE28837;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Discosoma sp. drFP583 (NFP-6) mutant protein, E5-NA.  
XX  
KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; mutain;  
KW



KW mutant.  
XX Discosoma sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Wild type Arg substituted with Ala"  
FT Misc-difference 5 /note= "Wild type Lys substituted with Glu"  
FT Misc-difference 9 /note= "Wild type Lys substituted with Thr"  
FT Misc-difference 105 /note= "Wild type Val substituted with Ala"  
FT Misc-difference 197 /note= "Wild type Ser substituted with Thr"  
FT  
XX  
PN WO200268459-A2.  
XX  
PD 06-SEP-2004.  
XX  
PF 20-FEB-2004; 2002WO-US005749.  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2004; 2001US-00006922.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX WPI; 2002-691654/74.  
XX  
DR N-PSDB; AAD46282.  
XX  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
PS Disclosure; Page; 80pp; English.  
XX  
XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as screens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. dfrp583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. dfrp583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX  
SQ Sequence 225 AA;  
  
Query Match 98.1%; Score 1191; DB 5; Length 225;  
Best Local Similarity 97.8%; Pred. No. 4e-125;  
Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSQKVIKEFMRPKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSENVITEFMRPKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFDGVTVTQDSSLQDGFY 120  
Db 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFDGVTVTQDSSLQDGFY 120  
QY 121 KVKFPGVNFPSDGPVMQKKTGWGASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFPGVNFPSDGPVMQKKTGWGASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKPVQLPGYVVVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKPVQLPGYVVVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

Db 181 YMAKPVQLPGYVVVDTKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
  
RESULT 75  
AAE17542  
ID AAE17542 standard; protein; 225 AA.  
XX  
AC AAE17542;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. non-aggregating mutant timer protein ESNA.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; mutant; ESNA; mutein.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "Wild type Arg substituted with Ala"  
FT Misc-difference 5 /note= "Wild type Lys substituted with Glu"  
FT Misc-difference 9 /note= "Wild type Lys substituted with Thr"  
FT  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Fradkov AF, Tersikh A;  
XX  
XX WPI; 2002-154595/20.  
XX  
DR N-PSDB; AAD28209.  
XX  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
XX  
PS Claim 5; Fig 3; 89pp; English.  
XX  
XX  
CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodelling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated structures, e.g.  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigations where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is Discosoma sp. ESNA







PN WO200268459-A2.  
XX  
XX 06-SEP-2002.  
PD  
XX  
XX 20-FEB-2002; 2002WO-US005749.  
PF  
XX  
XX 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
PR  
XX  
XX (CLON-) CLONTECH LAB INC.  
PA  
XX  
XX Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;  
XX  
XX WPI; 2002-691654/74.  
DR  
DR N-PSDB; AAD46281.  
XX  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
PS Disclosure; Page; 80pp; English.  
XX  
XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as screens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant protein of the  
CC invention. Note: This sequence is not shown in the specification, but is  
CC derived from Discosoma sp. drFP583 (NFP-6) wild-type protein shown as SEQ  
CC ID NO:8 (AAE28833) in page 70-71 of the specification  
XX  
XX  
SQ Sequence 225 AA;

Query Match	97.7%	Score 1186;	DB 5;	Length 225;
Best Local Similarity	97.3%	Pred. No. 1.5e-124;		
Matches 219; Conservative	2;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1	MSSSKAVIKEFMRFKVRMEGTVNGHEFEITEGEGEGRPYEGHNTVKLVTKGGLPFAWDI	60
		:	
Db	1	MASSENVITTEFMRFKVRMEGTVNGHEFEITEGEGEGRPYEGHNTVKLVTKGGLPFAWDI	60
Qy	61	LSPQFOYGSKVYVVKHPADIPDYKKLSPEGEFKMERVMNFEDGCVTVTQDSSLQDGCFIY	120
Db	61	LSPQFOYGSKVYVVKHPADIPDYKKLSPEGEFKMERVMNFEDGCVATVTVQDSSLQDGCFIY	120
Qy	121	KVKEFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGYLKGBIHKALKLKDGGHYLVEFKSI	180
Db	121	KVKEFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGYLKGETHKALKLKDGGHYLVEFKSI	180
Qy	181	YMAKKPVQLPGYIYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL	225
		:	
Db	181	YMAKKPVQLPGYIYVDAKLDITSHNEDYTIVEQYERTEGRHHLFL	225

RESULT 78	
ADH10227	
ID	ADH10227 standard; protein; 225 AA.

AC	ADH10227;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	Red fluorescent protein variant (DsRed).
XX	
KW	Protein immobilisation; red fluorescent protein; DsRed
XX	
OS	Synthetic.
XX	

PN EP1367121-A2.  
XX  
PD 03-DEC-2003.  
XX  
PF 23-MAY-2003; 2003EP-00253252.  
XX  
PR 23-MAY-2002; 2002JP-00148950.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX  
PI Iwakura M, Hirota K;  
XX  
DR WPI; 2004-045224/05.  
XX  
PT New carrier, useful for immobilizing polypeptides via the carboxyl group  
PT of the C-terminus, comprises a protein immobilized adsorptively, and  
PT having a sulfhydryl group.  
XX  
PS Example; SEQ ID NO 7; 31pp; English.  
XX  
CC The invention relates to a carrier comprising a protein immobilized  
CC adsorptively, where the protein has a sulfhydryl group. The carrier is  
CC useful for immobilizing polypeptides via the carboxyl group of the C-  
CC terminus. The present sequence represents a Red fluorescent protein  
CC variant (DsRed) sequence.  
XX  
SQ Sequence 225 AA;

Query Match	97.7%;	Score 1186;	DB 8;	Length 225;
Best Local Similarity	97.3%;	Pred. No. 1.5e-124;		
Matches 219;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	MRSSKAVIKEFMRFKVRMEGIYVNGHFEI	EGEGEGRP	YEGHNIVTLKVT	KGGLPF	AMDI	60
		:					
Db	1	MASSENVI	TEFMRFKVRMEGI	VNGHFEI	EGEGEGRP	YEGHNIVTLKVT	KGGLPF
							AMDI 60
QY	61	LSPQFOYGS	KVYVKHPADIP	PKKLSF	PPEGFKMERVMN	FEDGVA	VTYQDSSLQDGC
							FIY 120
Db	61	LSPQFOYGS	KVYVKHPADIP	PKKLSF	PPEGFKMERVMN	FEDGVA	VTYQDSSLQDGC
							FIY 120
QY	121	KYKFIGVNP	PSDGPVMQKKT	MGWEASTERL	YPRDGLKGEI	HKALKLKD	GHYLVEFKSI
							180
Db	121	KYKFIGVNP	PSDGPVMQKKT	MGWEASTERL	YPRDGLKGEI	HKALKLKD	GHYLVEFKSI
							180
QY	181	YMAKKPVQL	PGYYYVDSKLD	ITSHNEDY	TI	VEQYERTEGR	HHLFL 225
			:				
Db	181	YMAKKPVQL	PGYYYVDSKLD	ITSHNEDY	TI	VEQYERTEGR	HHLFL 225

RESULT 79  
ADT77526  
ID ADT77526 standard; protein; 225 AA.

DT	13-JAN-2005	(first entry)
XX		
DE	Protein SEQ ID NO:2	used in protein array.
XX		
KW	Protein array; planar	substrate; immobilisation; detection;
KW	target substance.	
XX		
OS	Unidentified.	
XX		
PN	WO2004090542-A1.	
XX		
PD	21-OCT-2004.	
XX		
PF	09-APR-2004; 2004WO-JP005150.	
XX		
PR	10-APR-2003; 2003JP-00106450.	
XX		
PA	(NAAD-) NAT INST	ADVANCED IND SCI & TECHNOLOGY.



XX  
PI Iwakura M, Hirota K;  
XX  
DR WPI; 2004-766496/75.  
XX  
PT Substrate for protein array, obtained by combining a polymer compound  
PT having primary amino groups in repeating units on surface of planar  
PT substrate.  
XX  
PS Example 1; SEQ ID NO 2; 32pp; Japanese.  
XX  
CC The invention relates to a substrate for a protein array, obtained by  
CC combining a polymer compound having primary amino groups in repeating  
CC units on a surface of a planar substrate. The invention also relates to a  
CC protein array formed by bonding the C-termini of proteins to the primary  
CC amino groups of the polymer compound on the substrate. The invention  
CC further relates to a method for the preparation of the protein array.  
CC Preferably, the planar substrate has water absorptivity, the polymer used  
CC is a polyallyl amine or a polylysine, and the immobilised proteins  
CC contain a linker peptide. The protein array of the invention enables the  
CC effective and highly sensitive detection of various target substances.  
CC Sequences ADT77525-ADT77526 represent proteins whose C-termini were  
CC joined to a negatively charged acidic linker peptide (ADT77529) to  
CC produce the sequences shown in ADT77527-ADT77528. Sequences ADT77527-  
CC ADT77528 were subsequently immobilised on a planar substrate via the  
CC linker peptide to form a protein array.  
XX  
SQ Sequence 225 AA;  
OY Query Match 97.7%; Score 1186; DB 8; Length 225;  
Best Local Similarity 97.3%; Pred. No. 1.5e-124;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MASSENVTTEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVAVTQTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVAVTQTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGETHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 80  
ADH10228  
ID ADH10228 standard; protein; 239 AA.  
XX  
AC ADH10228;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Immobilised red fluorescent protein variant (DsRed).  
XX  
KW Protein immobilisation; red fluorescent protein; DsRed.  
XX  
OS Synthetic.  
XX  
PN EP1367121-A2.  
XX  
PD 03-DEC-2003.  
XX  
PF 23-MAY-2003; 2003EP-00253252.  
XX  
PR 23-MAY-2002; 2002JP-00148950.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX  
PI Iwakura M, Hirota K;  
XX  
DR WPI; 2004-045224/05.  
XX  
PT New carrier, useful for immobilizing polypeptides via the carboxyl group  
PT of the C-terminus, comprises a protein immobilized adsorptively, and  
PT having a sulhydryl group.  
XX  
PS Example 3; SEQ ID NO 8; 31pp; English.  
XX  
CC The invention relates to a carrier comprising a protein immobilized  
CC adsorptively, where the protein has a sulhydryl group. The carrier is  
CC useful for immobilizing polypeptides via the carboxyl group of the C-  
CC terminus. The present sequence represents an immobilised red fluorescent  
CC protein variant (DsRed) sequence.  
XX  
SQ Sequence 239 AA;  
OY Query Match 97.7%; Score 1186; DB 8; Length 239;  
Best Local Similarity 97.3%; Pred. No. 1.6e-124;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
OY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
Db 1 MASSENVTTEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTGGPLPFAWDI 60  
OY 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVAVTQTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDEGVAVTQTQDSSLQDGCFTY 120  
OY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGETHKALKLKDGGHYLVEFKSI 180  
OY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 81  
ADT77528  
ID ADT77528 standard; protein; 239 AA.  
XX  
AC ADT77528;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Protein/linker peptide SEQ ID NO:4 used in protein array.  
XX  
KW Protein array; planar substrate; immobilisation; detection;  
KW target substance; linker peptide.  
XX  
OS Synthetic.  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..225  
FT Peptide /note= "Corresponds to SEQ ID NO:2"  
FT 226..239  
FT /note= "Corresponds to the linker peptide shown on page  
FT 12 of the specification"  
XX  
PN WO2004090542-A1.  
XX  
PD 21-OCT-2004.  
XX  
PF 09-APR-2004; 2004WO-JP005150.  
XX  
PR 10-APR-2003; 2003JP-00106450.  
XX  
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.



PI Iwakura M, Hirota K;  
XX  
DR WPI; 2004-766496/75.  
XX  
PT Substrate for protein array, obtained by combining a polymer compound  
PT having primary amino groups in repeating units on surface of planar  
PT substrate.  
XX  
PS Example 1; SEQ ID NO 4; 32pp; Japanese.  
XX  
CC The invention relates to a substrate for a protein array, obtained by  
CC combining a polymer compound having primary amino groups in repeating  
CC units on a surface of a planar substrate. The invention also relates to a  
CC protein array formed by bonding the C-termini of proteins to the primary  
CC amino groups of the polymer compound on the substrate. The invention  
CC further relates to a method for the preparation of the protein array.  
CC Preferably, the planar substrate has water absorptivity, the polymer used  
CC is a polyallyl amine or a polylysine, and the immobilised proteins  
CC contain a linker peptide. The protein array of the invention enables the  
CC effective and highly sensitive detection of various target substances.  
CC Sequences ADT7527-ADT7528 represent proteins comprising a C-terminal  
CC negatively charged acidic linker peptide (ADT7529) which were  
CC immobilised on a planar substrate via the linker peptide to form a  
CC protein array.  
XX  
SQ Sequence 239 AA;  
Query Match 97.7%; Score 1186; DB 8; Length 239;  
Best Local Similarity 97.3%; Pred. No. 1.6e-124;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTKGGPPLPFAWDI 60  
Db 1 MASSENVTTEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTKGGPPLPFAWDI 60  
QY 61 LSPQFOYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNPFEDGGVATVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGETHKALKKDGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYVVYVDAKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
RESULT 82  
ABP56686  
ID ABP56686 standard; protein; 240 AA.  
XX  
AC ABP56686;  
XX  
DT 25-MAR-2003 (first entry)  
XX  
DE Discosoma red fluorescent protein variant R18H,N24S,F125L and M183K.  
XX  
KM Mammalian codon optimised Discosoma red fluorescent protein; variant;  
KM Discosoma; red fluorescent protein; directed evolution; mutant.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
PN WO200294992-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 20-MAY-2002; 2002WO-US015968.  
XX  
PR 18-MAY-2001; 2001US-0291871P.  
XX  
PA (RIGE-) RIGEL PHARM INC.

XX  
PI Peelle B;  
XX  
DR WPI; 2003-120798/11.  
XX  
PT New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.  
XX  
PS Example 1; Page; 22pp; English.  
XX  
CC The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (I); (2)  
CC an isolated nucleic acid encoding (I); (3) a vector comprising the  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (I); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence represents a mammalian  
CC codon optimised Discosoma red fluorescent protein variant from the  
CC present invention. N.B. The present sequence is not given in the  
CC specification, but is derived from the protein in Fig 1 as stated in Fig  
XX  
SQ Sequence 240 AA;  
Query Match 97.4%; Score 1183; DB 6; Length 240;  
Best Local Similarity 97.3%; Pred. No. 3.5e-124;  
Matches 219; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGRPRYEGHNTVKLKVTKGGPPLPFAWDI 60  
Db 2 VRSSKNVIKEFMRFKVRMEGTVSGHEFEIEGEGRPRYEGHNTVKLKVTKGGPPLPFAWDI 61  
QY 61 LSPQFOYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNPFEDGGVVTVTQDSSLQDGCFTY 120  
Db 62 LSPQFOYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNPFEDGGVVTVTQDSSLQDGCFTY 121  
QY 121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEKSI 180  
Db 122 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGVLKGEIHKALKKDGGHYLVFEKSI 181  
QY 181 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 182 YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226  
RESULT 83  
ADL26532  
ID ADL26532 standard; protein; 226 AA.  
XX  
AC ADL26532;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma fluorescent protein dsFP583.  
XX  
KM fluorescent protein; fluorescent marker; gene expression;  
KM protein localization; FP.  
XX  
OS Discosoma sp.



XX PN WO2004016648-A1.  
XX PD 26-FEB-2004.  
XX PF 19-JUL-2003; 2003WO-DE002436.  
XX PR 19-JUL-2002; 2002DE-01033082.  
XX PA (AMAX-) AMAXA GMBH.  
XX PI Altrogge L, Males T;  
XX DR WPI; 2004-238793/22.  
XX PT New fluorescent proteins, useful as pharmaceuticals, biological markers  
PT and for optical information storage, comprise short-lived but restorable  
PT fluorescence.  
XX PS Example 1; Fig 1; 55bp; German.  
XX CC The present invention provides protein and coding sequences of novel  
CC fluorescent proteins. These are useful in pharmaceutical compositions and  
CC as fluorescent markers for gene expression and protein localization in  
CC cellular, developmental or molecular biology, particularly for measuring  
CC time-dependent cellular process such as protein diffusion and transport,  
CC also for optical information storage. The present sequence is a Discosoma  
CC fluorescent protein.  
XX SQ Sequence 226 AA;  
  
Query Match 96.8%; Score 1175.5; DB 8; Length 226;  
Best Local Similarity 97.8%; Pred. No. 2.2e-123;  
Matches 221; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRBSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
  
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180  
Db 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKST 180  
  
QY 181 YMAKK-PVOLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKAPVOLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226  
  
RESULT 84  
ADC24129 standard; protein; 225 AA.  
XX AC ADC24129;  
XX DT 18-DEC-2003 (first entry)  
XX DE Discosoma red fluorescent protein variant fast T1.  
XX KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
KW protein phosphatase; ion indicator; mutant; mutein; fast T1.  
XX OS Synthetic.  
XX OS Discosoma.  
XX PN US2003059835-A1.  
XX PD 27-MAR-2003.

XX PF 10-APR-2002; 2002US-00121258.  
XX PR 26-FEB-2001; 2001US-00794308.  
XX PR 24-MAY-2001; 2001US-00866538.  
XX PA (TSIE/) TSIE N Y.  
XX PA (CAMP/) CAMPBELL R E.  
XX PI Tsien RY, Campbell RE;  
XX DR WPI; 2003-743764/70.  
XX DR N-PSDB; ADC24128.  
XX PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
XX PS Example 1; SEQ ID NO 4; 67pp; English.  
XX CC The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma red fluorescent protein variant fast T1.  
XX SQ Sequence 225 AA;  
  
Query Match 95.6%; Score 1160; DB 7; Length 225;  
Best Local Similarity 96.0%; Pred. No. 1.2e-121;  
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSEDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTAKLKVTKGGLPFAWDI 60  
  
QY 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGSFTY 120  
  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180  
Db 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFKSI 180  
  
QY 181 YMAKKPVOLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVOLPGYVVVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225



RESULT 85	
ADL46206	
ID	ADL46206 standard; protein; 225 AA.
XX	
AC	ADL46206;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Discosoma red fluorescent protein (DsRed) variant T1 protein.
XX	
KW	red fluorescent protein; DsRed; fluorescence; red wavelength;
KM	oligomerization; tetramerization; immunoassay; hybridization assay.
XX	
OS	Discosoma sp.
XX	
PN	WO2003086446-A1.
XX	
PD	23-OCT-2003.
XX	
PF	09-APR-2003; 2003WO-US010879.
XX	
PR	10-APR-2002; 2002US-00121258.
PR	29-JUL-2002; 2002US-00209208.
XX	
PA	(REGC ) UNIV CALIFORNIA.
XX	
PI	Tsien RY, Campbell RE, Baird GS;
XX	
DR	WPI; 2003-845265/78.
DR	N-PSDB; ADL46207.
XX	
PT	New monomeric and dimeric Anthozoan fluorescent protein variants with
PT	reduced propensity to oligomerize, and encoding polynucleotides, useful
PT	in molecular biology, e.g. in immunoassays or in tracking protein
PT	movement in cells.
XX	
PS	Disclosure; SEQ ID NO 4; 166pp; English.
XX	
CC	The invention relates to a polynucleotide sequence encoding a Discosoma
CC	red fluorescent protein (DsRed) variant having a reduced propensity to
CC	oligomerize. The protein variant comprises one or more amino acid
CC	substitutions at the AB and/or AC interface(s) of the wild-type DsRed
CC	sequence, where the substitutions result in reduced propensity of the
CC	DsRed variant to form tetramers and where the variant displays detectable
CC	fluorescence of at least one red wavelength. The composition and methods
CC	are useful in producing red fluorescent proteins having reduced
CC	propensity for oligomerization, especially tetramerization. The protein
CC	may be used in molecular biology and in other scientific applications,
CC	such as in immunoassays or hybridization assays, or in tracking the
CC	movement of proteins in cells. This sequence corresponds to the DsRed
CC	variant T1 protein.
XX	
SQ	Sequence 225 AA;
Query Match	95.6%; Score 1160; DB 7; Length 225;
Best Local Similarity	96.0%; Pred. No. 1.2e-121;
Matches 216; Conservative	3; Mismatches 6; Indels 0; Gaps 0;
QY	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db	1 MASSEDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQAKLKVTKGGLPFAWDI 60
QY	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGCFTY 120
Db	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGSFTY 120
QY	121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFSI 180
Db	121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFSI 180
QY	181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db	181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERAEGRHHLFL 225

RESULT 86	
ADL46226	
ID	ADL46226 standard; protein; 225 AA.
XX	
AC	ADL46226;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Discosoma red fluorescent protein (DsRed) T1 protein with I125R mutation.
XX	
KW	red fluorescent protein; DsRed; fluorescence; red wavelength;
KM	oligomerization; tetramerization; immunoassay; hybridization assay.
XX	
OS	Discosoma sp.
XX	
PN	WO2003086446-A1.
XX	
PD	23-OCT-2003.
XX	
PF	09-APR-2003; 2003WO-US010879.
XX	
PR	10-APR-2002; 2002US-00121258.
PR	29-JUL-2002; 2002US-00209208.
XX	
PA	(REGC ) UNIV CALIFORNIA.
XX	
PI	Tsien RY, Campbell RE, Baird GS;
XX	
DR	WPI; 2003-845265/78.
XX	
PT	New monomeric and dimeric Anthozoan fluorescent protein variants with
PT	reduced propensity to oligomerize, and encoding polynucleotides, useful
PT	in molecular biology, e.g. in immunoassays or in tracking protein
PT	movement in cells.
XX	
PS	Disclosure; SEQ ID NO 24; 166pp; English.
XX	
CC	The invention relates to a polynucleotide sequence encoding a Discosoma
CC	red fluorescent protein (DsRed) variant having a reduced propensity to
CC	oligomerize. The protein variant comprises one or more amino acid
CC	substitutions at the AB and/or AC interface(s) of the wild-type DsRed
CC	sequence, where the substitutions result in reduced propensity of the
CC	DsRed variant to form tetramers and where the variant displays detectable
CC	fluorescence of at least one red wavelength. The composition and methods
CC	are useful in producing red fluorescent proteins having reduced
CC	propensity for oligomerization, especially tetramerization. The protein
CC	may be used in molecular biology and in other scientific applications,
CC	such as in immunoassays or hybridization assays, or in tracking the
CC	movement of proteins in cells. This sequence corresponds to the DsRed
CC	variant T1 protein with a I125R mutation.
XX	
SQ	Sequence 225 AA;
Query Match	95.0%; Score 1153; DB 7; Length 225;
Best Local Similarity	95.6%; Pred. No. 7.6e-121;
Matches 215; Conservative	3; Mismatches 7; Indels 0; Gaps 0;
QY	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
Db	1 MASSEDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQAKLKVTKGGLPFAWDI 60
QY	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGCFTY 120
Db	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKWERNVNFEDGVTVTQDSSLQDGSFTY 120
QY	121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFSI 180
Db	121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVFEFSI 180
QY	181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225



Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERABGRHHLFL 225

RESULT 87

ADM13537

ID ADM13537 standard; protein; 225 AA.

XX

AC ADM13537;

XX

DT 15-JUL-2004 (first entry)

XX

DE Superfolding coral DsRed protein #1.

XX

KW GFP, superfolder mutant; mutein; green fluorescent protein;

KW directed-evolution; folding interference domain; reporter protein;

KW mutant; DsRed; coral.

XX

OS Discosoma sp.

OS Synthetic.

XX

PN US2004078148-A1.

XX

PD 22-APR-2004.

XX

PF 24-APR-2003; 2003US-00423688.

XX

PR 24-APR-2002; 2002US-00132067.

XX

PA (WALD/) WALDO G S.

XX

PI Waldo GS;

XX

DR WPI; 2004-340059/31.

DR N-PSDB; ADM13536.

XX

PT New modified green fluorescent protein, useful as a reporter in expression studies.

PS Claim 18; SEQ ID NO 41; 46pp; English.

XX

XX The invention relates to a green fluorescent protein (GFP) that comprises at least 80% identity to the GFP superfolding mutant appearing as ADM13535 that has at least one amino acid substitution selected from e.g., a substitution at position 30 that is an arginine or a conservative variant of arginine and measurable fluorescence activity. Also included are directed-evolution for generating an enhanced folding variant of a target polypeptide (comprising mutating a polynucleotide encoding polypeptide of interest to generate a library of mutated polynucleotides, linking mutated polynucleotides to polynucleotide encoding folding interference domain to form fusion protein (FP) constructs, expressing FP, and selecting FP that display optimal folding activity in relation to FP comprising wild-type polypeptide and folding interference domain, thus identifying polypeptide with enhanced folding activity), an enhanced folding variant of a fluorescent protein generated by the method, an enhanced folding variant of a chromophoric protein generated by the method, enhancing folding of a polypeptide comprising multiple domains (involving joining a first domain of the polypeptide to a poorly folding domain, to form a fusion protein, mutating the first domain, detecting an increase in the amount of activity generated by a first mutated fusion protein in comparison to a fusion protein comprising a wild-type first domain and the poorly folding polypeptide domain, thus identifying a first domain with enhanced folding, joining a second domain of the polypeptide to the first mutated fusion protein to form a second fusion protein, mutating the second domain, and detecting an increase in the amount of activity generated by a second mutated fusion protein in comparison to a fusion protein comprising the wild-type second domain and the first mutated fusion protein, thus identifying a target polypeptide with multiple domains that have enhanced folding), an isolated nucleic acid encoding the GFP, an expression vector comprising the nucleic acid and a host cell comprising the vector. The DsRed protein from coral is also considered for mutation using the above method. The enhanced folding variants are useful as reporter proteins to express the report level of a protein. The method is useful for directed-evolution of generating

CC enhanced folding variant of target polypeptide and for improving folding

CC and solubility of a target protein. The present sequence represents a

CC coral DsRed enhanced folding variant. NOTE: The authors refer to SEQ ID

CC 38-42 as SEQ ID 1-5 throughout the specification and never refer to SEQ

CC ID 1-37 by number, it is obvious that SEQ ID 38-42 are meant to be SEQ ID

CC 1-5.

XX

SQ Sequence 225 AA;

Query Match 93.2%; Score 1132; DB 8; Length 225;

Best Local Similarity 92.4%; Pred. No. 1.8e-118;

Matches 208; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MRSSKNVIEKPMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDI 60

Db 1 MESSEDDYIKBFMRFKVHMEGSVNGHEFEIEGEGRPYEGTQNVKLVKTKGGLPFAMDI 60

QY 61 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKWERVNNFEDGVTVTQDSSLQDGCFTY 120

Db 61 LSPQFYGSKVYVKHPADIPDYKKLSFPEGFKWERVNNFEDGVTVTQDSSLQDGCFTY 120

QY 121 KYKFIGVNFPSDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKKDGGHYLVFEFSI 180

Db 121 KYKFIGVNFPSDGPVMQKKTWGEASTERLYPRDGLKGEIHKALKKDGGHYLVDIKSI 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERABGRHHLFL 225

RESULT 88

ADC24131

ID ADC24131 standard; protein; 226 AA.

XX

AC ADC24131;

XX

DT 18-DEC-2003 (first entry)

XX

DE Discosoma red fluorescent protein variant dimer2.

XX

KW Discosoma red fluorescent protein; DsRed; AB interface; AC interface;

KW fluorescent protein variant; transcription induction detection;

KW fluorescence energy resonance transfer; FRET; protein kinase;

XX protein phosphatase; ion indicator; mutant; dimer2; mutein.

OS Synthetic.

OS Discosoma.

XX

PN US2003059835-A1.

XX

PD 27-MAR-2003.

XX

PF 10-APR-2002; 2002US-00121258.

XX

PR 26-FEB-2001; 2001US-00794308.

PR 24-MAY-2001; 2001US-00866538.

XX

PA (TSIE/) TSIEH R Y.

PA (CAMP/) CAMPBELL R E.

XX

PI Tsien RY, Campbell RE;

XX

DR WPI; 2003-743764/70.

DR N-PSDB; ADC24130.

XX

PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein variant having a reduced propensity to oligomerize, useful for detecting transcriptional activity.

XX

PS Claim 12; SEQ ID NO 6; 67pp; English.

XX

CC The invention describes a polynucleotide sequence (I) encoding a

CC Discosoma red fluorescent protein (DsRed) variant having a reduced



CC propensity to oligomerise, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or controlling that regulates the activity  
CC of an expression control sequence. This is the amino acid sequence of  
CC Discosoma red fluorescent protein variant dimer2.

XX Sequence 226 AA;

Query Match 92.3%; Score 1121; DB 7; Length 226;

Best Local Similarity 92.8%; Pred. No. 3.1e-117; Indels 0; Gaps 0;  
Matches 207; Conservative 7; Mismatches 9;

QY 3 SSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDILS 62  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
4 SSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQAKLKVTKGGLPFAMDILS 63  
QY 63 PQFYGSKVYVKGHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFIYKV 122  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 64 PQFYGSKAVYVKGHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGLIYKV 123  
QY 123 KFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSIYM 182  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 124 KFRGTNFPDPGPVMQKTMGWEASTERLYPRDGLKGEIHQALKDGGHYLVEFKTIYM 183  
QY 183 AKKPVLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 184 AKKPVLPGYYVYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 226

RESULT 89

ADL46208

ID ADL46208 standard; protein; 226 AA.

XX ADL46208;

DT 20-MAY-2004 (first entry)

XX Discosoma red fluorescent protein (DsRed) variant dimer2 protein.

DE red fluorescent protein; DsRed; fluorescence; red wavelength;

KW oligomerization; tetramerization; immunoassay; hybridization assay.

XX Discosoma sp.

XX WO2003086446-A1.

XX 23-OCT-2003.

PF 09-APR-2003; 2003WO-US010879.

XX 10-APR-2002; 2002US-00121258.

PR 29-JUL-2002; 2002US-00209208.  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Tsien RY, Campbell RE, Baird GS;  
PI  
XX WPI; 2003-845265/78.  
DR N-PSDB; ADL46209.

XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.

PS Disclosure; SEQ ID NO 6; 166pp; English.

XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC variant dimer2 protein.

XX Sequence 226 AA;

Query Match 92.3%; Score 1121; DB 7; Length 226;

Best Local Similarity 92.8%; Pred. No. 3.1e-117; Indels 0; Gaps 0;  
Matches 207; Conservative 7; Mismatches 9;

QY 3 SSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAMDILS 62  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 4 SSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQAKLKVTKGGLPFAMDILS 63  
QY 63 PQFYGSKVYVKGHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGCFIYKV 122  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 64 PQFYGSKAVYVKGHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGLIYKV 123  
QY 123 KFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKDGGHYLVEFKSIYM 182  
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 124 KFRGTNFPDPGPVMQKTMGWEASTERLYPRDGLKGEIHQALKDGGHYLVEFKTIYM 183  
QY 183 AKKPVLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
DB 184 AKKPVLPGYYVYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 226

RESULT 90

AA99838

ID AA99838 standard; protein; 225 AA.

XX AA99838;

DT 19-SEP-2000 (first entry)

XX Humanised anthozoa novel fluorescent protein hybrid drFP583/dmFP592.

XX Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;

KW fluorescent labeling; hybrid.

XX Discosoma sp; "red".

XX Synthetic.

XX WO200034326-A1.

XX 15-JUN-2000.



PF 10-DEC-1999; 99WO-US029473.  
XX 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
DR WPI; 2000-423381/36.  
DR N-PSDB; AAA48748.  
XX  
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
PS Claim 20; Page 78; 86pp; English.  
XX  
CC The present sequence is humanised drFP583/dmFP592. Amplified wild-type  
CC coding region fragments from drFP583 and dmFP592 were digested with  
CC restriction endonucleases and combined together to produce a hybrid  
CC construct. drFP583 is a full-length cDNA encoding a novel fluorescent  
CC protein (nFP) from Discosoma sp. "red", a non-bioluminescent species of  
CC the Class Anthozoa. The hybrid sequence was changed to optimise the  
CC codons for expression of the protein in mammalian cells. Fluorescent  
CC proteins can be used in fluorescent labeling, a useful tool for marking a  
CC protein, cell or organism of interest. Unlike other markers used in  
CC protein labeling, such as beta-galactosidase and luciferase, fluorescent  
CC proteins do not require an exogenous cofactor or substrate. Methods  
CC involving fluorescent proteins are also less laborious and less difficult  
CC to control than the traditional methods of fluorescent labeling, where a  
CC protein of interest is purified and then covalently conjugated to a  
CC fluorophore derivative. Novel fluorescent proteins isolated from species  
CC of the Class Anthozoa can be used as markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations compared to prior art fluorescent  
CC proteins such as green fluorescent protein  
XX  
SQ Sequence 225 AA;

Query Match 92.2%; Score 1119; DB 3; Length 225;  
Best Local Similarity 91.6%; Pred. No. 5.1e-117;  
Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MSCSKNVIKEFMRFKVRMEGTVNGHEFEIKGEGGRPYEGHCSVKLMVTKGGLPFAFDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 91  
AAE28835  
ID AAE28835 standard; protein; 225 AA.  
XX  
AC AAE28835;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Discosoma sp. drFP583/dmFP592 (6/9Q) protein.  
XX  
KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;

KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; dmFP592 protein.  
XX  
OS Discosoma sp.  
XX  
PN WO200268459-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 20-FEB-2002; 2002WO-US005749.  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;  
XX  
DR WPI; 2002-691654/74.  
DR N-PSDB; AAD46280.  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
XX applications.  
XX  
PS Disclosure; Page 72-73; 80pp; English.  
XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583/dmFP592 (6/9Q) protein of the invention  
XX  
SQ Sequence 225 AA;

Query Match 92.2%; Score 1119; DB 5; Length 225;  
Best Local Similarity 91.6%; Pred. No. 5.1e-117;  
Matches 206; Conservative 12; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MSCSKNVIKEFMRFKVRMEGTVNGHEFEIKGEGGRPYEGHCSVKLMVTKGGLPFAFDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFBEDGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
Db 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKDKGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 92  
AAI99839  
ID AAI99839 standard; protein; 225 AA.  
XX  
AC AAI99839;  
XX  
DT 19-SEP-2000 (first entry)  
XX  
DE Humanised Anthozoa novel fluorescent protein hybrid drFP583/dmFP592-2G.  
XX



KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
 KW fluorescent labeling; hybrid; mutant; mutein.  
 XX  
 OS Discosoma sp; "red".  
 OS Synthetic.  
 XX  
 PN WO200034326-A1.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US029473.  
 XX  
 PR 11-DEC-1998; 98US-00210330.  
 PR 14-OCT-1999; 99US-00418529.  
 XX  
 PA (CLON-) CLONTECH LAB INC.  
 XX  
 PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
 PI Ding L;  
 XX  
 DR WPI; 2000-423381/36.  
 DR N-PSDB; AAA48749.  
 XX  
 PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
 PT useful for fluorescent labeling and as markers.  
 XX  
 PS Claim 6; Page 79; 86pp; English.

The present sequence is derived from humanised hybrid protein drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel fluorescent protein (nFP) from *Dicostoma* sp. "red", a non-bioluminescent species of the Class Anthozoa. The hybrid protein sequence was changed to optimise the codons for expression of the protein in mammalian cells. The present sequence contains two substitutions, K15Q and T217S, which were incorporated by random mutagenesis. Fluorescent proteins can be used in fluorescent labeling, a useful tool for marking a protein, cell or organism of interest. Unlike other markers used in protein labeling, such as beta-galactosidase and luciferase, fluorescent proteins do not require an exogenous cofactor or substrate. Methods involving fluorescent proteins are also less laborious and less difficult to control than the traditional methods of fluorescent labeling, where a protein of interest is purified and then covalently conjugated to a fluorophore derivative. Novel fluorescent proteins isolated from species of the Class Anthozoa can be used as markers for gene expression and protein localization studies, and in fluorescence resonance energy transfer (FRET) reactions. They may have improved properties and better suitability for larger excitations compared to prior art fluorescent proteins such as green fluorescent protein

Sequence 225 AA;

Query Match	91.5%	Score 1111;	DB 3;	Length 225;
Best Local Similarity	90.7%;	Pred. No. 4.1e-116;		
Matches 204; Conservative	14;	Mismatches 7;	Indels 0;	Gaps 0;

```
QY      1 MSSKNVKEFMRFKVRMEGTWNGHFEI BGE GEGRPYE GHN TVKLVT KGGPLPFAWDI    60
        | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db      1 MSCSKVIKEFMRFOVRMEGTWNGHFEI KGEGEGRPYEGHCSVKLMTWKGP L PFAFDI    60
```

QY	61	LS	PQ	FQ	YGS	KV	VK	HP	AD	IP	DY	KCL	S	P	E	G	F	K	N	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	L	D	G	C	F	I	Y	120
Db	61	LS	PQ	FQ	YGS	KV	VK	HP	AD	IP	DY	KCL	S	P	E	G	F	K	N	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	L	D	G	C	F	I	Y	120

Dy            121 KVFFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGVLKEIHKALKKDGGHYLVEEKS I 180  
       :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db          121 EVKFIGNPSPDGPVMQRTRTGWEASSERLYPRDGVLKGDIMHALLRLEGGGHYLV EKS I 180

QY	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHLL	FL	225
			:	
Db	181	YMAKKPVQLPGYYVDSKLDITSHNEDYTIVEQYERSEGRHLL	FL	225

## RESULT 93

AAE28928  
ID AAE28928 standard; protein; 225 AA.

AC AAE28928;

DT 27-DEC-2002 (first entry)

Discovoma sp. dRFP583/dmFP592 (6/9Q) mutant protein (K15Q+T217S).

AA Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
 KW fluorescence activated cell sorting application; fluorescent timer;  
 KW biosensor; fluorescence resonance energy transfer application; FRET;  
 KW colouring agent; recombinant DNA application; analyte detection assay;  
 KW sunscreen; second messenger detector; drFP583 protein; dmFP592 protein;  
 KW mutant; mutuin.

05 *Discosoma* sp.  
05 Synthetic.

AA	Key	Location/Qualifiers
FH		

```

F1 MISC-DIFFERENCE 13
FT /note= "Wild-type Lys substituted with Gln"

```

FT M18c-difference 21, /note= "Wild-type Thr substituted with Ser"

PN WO200268459-A2

06-SEP-2002.  
PD

20-FEB-2002; 2002WO-US005749.

21-FEB-2001; 2001US-0270983P.

04-DEC-2001; 2001US-00006922.

PA (CLON-) CLONTECH LAB INC.

AA  
PI Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;

WPI: 2002-691654/74.

xx New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Cnidarian chromo- or fluorescent protein or mutant for  
PT an aggregate detection assay or fluorescence activated cell sorting  
PT applications.

aa Disclosure; Page; 80pp; English.

xx The invention relates to nucleic acid molecules encoding non-aggregating  
cc chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
cc useful in analyte detection assays, as colouring agents, as markers in  
cc recombinant DNA applications, as sunscreens or filters, in fluorescence  
cc resonance energy transfer (FRET) applications, as biosensors in  
cc prokaryotic and eukaryotic cells, in screening assays, as second  
cc messenger detectors, in fluorescence activated cell sorting applications,  
cc in protease cleavage assays or as fluorescent timers. The present  
cc sequence is *DiscoSoma* sp. drFP583/dmFP592 (6/90) mutant protein of the  
cc invention. Note: This sequence is not shown in the specification, but is  
cc derived from *DiscoSoma* sp. drFP583/dmFP592 (6/90) wild-type protein shown  
cc as SEQ ID NO:12 (AAE28835) in page 72-73 of the specification

5Q Sequence 225 AA;

Query Match	91.5%;	Score 1111;	DB 5;	Length 225;
Best Local Similarity	90.7%;	Pred. No. 4.1e-116;		
Matches 204; Conservative	14;	Mismatches 7;	Indels 0;	Gaps 0;

```
QY      1 MESSKNVKEFMRFKVRMEGTGNCHPEIEGEGRPYEGHNTVKLKTKGGLPFAWDI   60  
        | ||||| :||||| :||||| :||||| :||||| :||||| :||||| :  
Db      1 MSCSKNVKEFMRFQVRMEGTGNCHPEIKGEGRPRYEGHC VKLMWTKGGLPFAADI   60
```

```
QY      61 LSPQFOYGSKVYVKHPADIPDYKLLSPPEGFKMERVNNFEDGCVTVTQDSSLQDGCFTY 120C
        |||||:|||||
Db      61 LSPQFOYGSKVYVKHPADIPDYKLLSPPEGFKMERVNNFEDGCVTVTQDSSLQDGCFTY 120C
```



QY	121	KVKFIGVNFPSDGPVMOKTGMWEASTERRLYPRDGLVKGELHKKLTKDGGHYLVEFKSI	180
Db	121	EVKFIGVNFPSDGPVMQRTRGWEASSERLYPRDGLVKGDIHMLRLREGGHYLVEFKSI	180
QY	181	YMAKPPVOLPGYTYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL	225
Db	181	YMAKPPVOLPGYTYVDSKLDITSHNEDYTYVEQYERSEGRHHLFL	225
RESULT 94			
ID	AAE28929	standard; protein; 223 AA.	
XX	AAE28929;		
DT	27-DEC-2002	(first entry)	
DE	Discosoma sp. drFP583/dmFP592 (6/9Q)	mutant protein (Sdel+Cdel+K5E+K9T).	
XX			
KW	Fluorescent protein; chromoprotein; protease cleavage assay; filter;		
KW	fluorescence activated cell sorting application; fluorescent timer;		
KW	biosensor; fluorescence resonance energy transfer application; FRET;		
KW	colouring agent; recombinant DNA application; analyte detection assay;		
KW	sunscreen; second messenger detector; drFP583 protein; dmFP592 protein;		
KW	mutant; mutein.		
OS	Discosoma sp.		
OS	Synthetic.		
FH	Key	Location/Qualifiers	
FT	Misc-difference	5	/note= "Wild-type Lys substituted with Glu"
FT	Misc-difference	9	/note= "Wild-type Lys substituted with Thr"
XX	WO200268459-A2.		
PD	06-SEP-2002.		
PF	20-FEB-2002; 2002WO-US005749.		
PR	21-FEB-2001; 2001US-0270983P.		
PR	04-DEC-2001; 2001US-00006922.		
PA	(CLON-) CLONTECH LAB INC.		
PI	Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;		
DR	WPI; 2002-691654/74.		
XX			
PT	New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant		
PT	of an aggregating Cnidarian chromo- or fluorescent protein or mutant for		
PT	analyte detection assays or fluorescence activated cell sorting		
PT	applications.		
XX			
PS	Disclosure; Page; 80pp; English.		
XX			
CC	The invention relates to nucleic acid molecules encoding non-aggregating		
CC	chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are		
CC	useful in analyte detection assays, as colouring agents, as markers in		
CC	recombinant DNA applications, as sunscreens or filters, in fluorescence		
CC	resonance energy transfer (FRET) applications, as biosensors in		
CC	prokaryotic and eukaryotic cells, in screening assays, as second		
CC	messenger detectors, in fluorescence activated cell sorting applications,		
CC	in protease cleavage assays or as fluorescent timers. The present		
CC	sequence is Discosoma sp. drFP583/dmFP592 (6/9Q) mutant protein of the		
CC	invention. Note: This sequence is not shown in the specification, but is		
CC	derived from Discosoma sp. drFP583/dmFP592 (6/9Q) wild-type protein shown		
CC	as SEQ ID NO:12 (AAE28835) in page 72-73 of the specification		
XX			
XX	Sequence 223AA;		

	Query Match	91.1%;	Score 1106;	DB 5;	Length 223;	
	Best Local Similarity	91.4%;	Pred. No. 1.5e-115;			
	Matches	203;	Conservative	13;	Mismatches	6; Indels 0; Gaps 0

  

QY	4	SKNVIKFPMRFKVMEGTVNGHFEIEGEGEGRPYEGHNTVKLKYTKGGPLPFAWDILSP	63
Dd	2	SENVTTFMRFKVMEGVNNGHEFEIKGEGERPRYEHCYSVKLMTYKGGLPFPAFDILSP	61
QY	64	QFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVWNFEDGGVTVTQDSSLDDGCFIYVKV	123
Dd	62	QFOYGSKVYVKHPADIPDYKKLSFPEGFKWERVWNFEDGGVTVTSQDSLDDGCFIYEVK	121
QY	124	FIVGNFPSDGPVMOKTGMWEASTERLYPRDVLKGEIHKALK.KDGGHYLVFEKSIYMA	183
Dd	122	FIVGNFPSDGPVMQRRTRGWEASSERLYPRDVLKGDIMALRLGCGHYLVFEKSIYMA	181
QY	184	KKPVOQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL	225
Dd	182	KKPVOQLPGYYVDISKLDITSHNEDYTIVEQYERTEGRHHLFL	223

  

RESULT 95	
AAY99840	ID AAY99840 standard; protein; 225 AA.
XX	AAY99840;
AC	19-SEP-2000 (first entry)
DT	
DE	Humanised Anthozoa novel fluorescent protein hybrid drFP583/dmFP592-Q3.
XX	
KW	Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;
KW	fluorescent labeling; hybrid; mutant; mutein.
XX	
OS	Discosoma sp; "red".
OS	Synthetic.
PN	WO200034326-A1.
PD	15-JUN-2000.
PF	10-DEC-1999; 99WO-US029473.
PR	11-DEC-1998; 98US-00210330.
PR	14-OCT-1999; 99US-00418529.
XX	(CLON-) CLONTech LAB INC.
PI	Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;
PI	Ding L;
DR	WPI; 2000-423381/36.
DR	N-PSDB; AAA48750.
PT	Novel fluorescent protein from non-bioluminescent Discosoma sp. red,
PT	useful for fluorescent labeling and as markers.
PS	Claim 20; Page 81-82; 86pp; English.

  

The present sequence is derived from humanised hybrid protein drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel fluorescent protein (nfp) from *Discosoma* sp. "red", a non-bioluminescent species of the Class Anthozoa. The hybrid sequence was changed to optimise the codons for expression of the protein in mammalian cells. The present sequence contains three substitutions, K15Q, K83M and T217S, which were incorporated by random mutagenesis. Fluorescent proteins can be used in fluorescent labeling, a useful tool for marking a protein, cell or organism of interest. Unlike other markers used in protein labeling, such as beta-galactosidase and luciferase, fluorescent proteins do not require an exogenous cofactor or substrate. Methods involving fluorescent proteins are also less laborious and less difficult to control than the traditional methods of fluorescent labeling, where a











PR 09-DEC-1999; 99US-00457898.  
PR 09-DEC-1999; 99US-00458144.  
PR 09-DEC-1999; 99US-00458477.  
PR 10-DEC-1999; 99WO-US029405.  
PR 14-JUN-2000; 2000US-0211607P.  
PR 14-JUN-2000; 2000US-0211609P.  
PR 14-JUN-2000; 2000US-0211626P.  
PR 14-JUN-2000; 2000US-0211627P.  
PR 14-JUN-2000; 2000US-0211687P.  
PR 14-JUN-2000; 2000US-0211766P.  
PR 14-JUN-2000; 2000US-0211880P.  
PR 14-JUN-2000; 2000US-0211888P.  
PR 14-JUN-2000; 2000US-0212070P.  
XX  
XX  
PA (CLON-) CLONTECH LAB INC.  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Terakikh A;  
XX  
XX  
DR WPI; 2001-266409/27.  
DR N-PSDB; AAD03618.  
XX  
XX  
PT An Anthozoa chromo- or fluorescent protein (P1) present in an environment  
PT other than its natural environment, useful as a label in analyte  
PT detection assays and as a selectable marker in recombinant DNA  
PT applications.  
XX  
XX  
PS Claim 8; Fig 9; 69pp; English.  
XX  
CC The present sequence is a Discosoma sp. magenta chromo/fluorescent  
CC protein, dmp592 (NFP-9). NFP-9 is present in an environment other than  
CC its natural environment and has an absorbance maximum ranging from 375nm  
CC to 775nm and more usually from 560nm to 590nm and emission maximum  
CC ranging from 395nm to 795nm and more usually from 580 to 610nm. The  
CC chromoproteins or fluorescent proteins are useful as labels in analyte  
CC detection assays, as selectable markers in recombinant DNA applications,  
CC as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator  
CC and as in vivo markers in animals. They are also useful in sunscreens and  
CC as selective filters. (Updated on 11-SEP-2003 to standardise OS field)  
XX  
XX  
SQ Sequence 230 AA;

Query Match 89.4%; Score 1085.5; DB 4; Length 230;  
Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MRSSKNVIKEFMRFKYRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 1 MSCSKNVIKEFMRFKYRMEGTVNGHEFEIKGEGRPYEGHCSVKLMVTKGGLPFAFDI 60  
QY 61 LSPQFOYGSKVYVKPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGSHYLVFPKSI 180  
DB 121 EVKFIGVNFPSDGPVMQKTRGWEASSERLYPRDGLKGDIMALRLBGGGHYLVFPKSI 180  
QY 181 YMAKKP-VQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMWKKPSVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTQGRHHPFI 226

RESULT 100  
ABP70038  
ID ABP70038 standard; protein; 230 AA.  
XX  
AC ABP70038;  
XX  
DT 06-AUG-2003 (revised)  
DT 22-JAN-2003 (first entry)  
XX  
DE Colour facilitating molecule (CFM) related sequence #SEQ ID 240.  
XX

KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;  
KW chromophore; biomatrix; transgenic animal; colouring agent;  
KW flower industry; expression marker; reporter molecule; photon trap;  
KW UV sink; sunscreen.  
XX  
OS Unidentified.  
XX  
XX WO200270703-A2.  
PN  
XX  
XX 12-SEP-2002.  
PD  
XX  
XX 01-MAR-2002; 2002WO-GB000928.  
PF  
XX  
XX 02-MAR-2001; 2001US-0273227P.  
PR 21-MAR-2001; 2001AU-00003874.  
PR 15-OCT-2001; 2001US-0329816P.  
XX  
XX (NUFA-) NUFARM LTD.  
PA (UYQU ) UNITV QUEENSLAND.  
PA (JONE/) JONES E L.  
XX  
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;  
PI Hoegh-Guldberg IO, Prescott M;  
XX  
XX WPI; 2002-740765/80.  
DR  
XX  
XX  
XX  
PT Novel color-facilitating molecule for producing a biomatrix, has a  
PT polypeptide which alone/along with molecules imparts altered visual  
PT characteristics to cells in the absence of excitation by extraneous non-  
PT white light.  
XX  
XX  
PS Example 19; Page 498; 510pp; English.

XX  
CC The invention relates to an isolated colour-facilitating molecule (CFM)  
CC comprising a polypeptide which, in a cell, alone or together with one or  
CC more other molecules imparts an altered visual characteristic to the cell  
CC when visualised by a human eye in the absence of excitation by extraneous  
CC non-white light or particle emission. CFMs are useful for producing a  
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or  
CC red coloured fleece. They are useful for producing coloured plant  
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other  
CC uses include transducing or intensifying an image, providing additional  
CC light for growing phototropic organisms e.g. algae and/or corals, for  
CC coating materials that experience UV damage e.g. plastics and car  
CC upholstery. CFMs are useful in the flower industry, in the development of  
CC new varieties of flowering plants. Other contemplated uses include,  
CC expression markers, general reporter molecules, photon traps, UV sinks or  
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental  
CC fungal species, and in fruits and vegetables to enhance their  
CC marketability. CFMs embedded in a gel matrix improve image quality in  
CC situations of distorted light spectra (biomatrix). The first all-protein  
CC chromophore to be isolated was Green Fluorescent protein (GFP). The  
CC sequences given in records ABP69924-ABP70048 represent CFM related amino  
CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)  
XX  
XX  
SQ Sequence 230 AA;

Query Match 89.4%; Score 1085.5; DB 5; Length 230;  
Best Local Similarity 88.1%; Pred. No. 3.1e-113;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;

QY 1 MRSSKNVIKEFMRFKYRMEGTVNGHEFEIEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60  
DB 1 MSCSKNVIKEFMRFKYRMEGTVNGHEFEIKGEGRPYEGHCSVKLMVTKGGLPFAFDI 60  
QY 61 LSPQFOYGSKVYVKPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFOYGSKVYVKPADIPDYKKLSFPEGFKMERVMNFDGCVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGSHYLVFPKSI 180  
DB 121 EVKFIGVNFPSDGPVMQKTRGWEASSERLYPRDGLKGDIMALRLBGGGHYLVFPKSI 180



```

Oy      181 YMAKKP-VQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :||| ||| ||| :
Db      181 YMVKKPSVOLPGYYVDISKLDMTSHNEDYTIVVEQYEKTQGRHHPFI 226

RESULT 101
ID      ABP96651 standard; protein; 230 AA.
XX      AC      ABP96651;
XX      DT      29-MAY-2003 (first entry)
DE      Red fluorescent protein (DsRed) protein SEQ ID NO:40.
KW      Chromosome-based platform; artificial chromosome; eukaryotic chromosome;
KM      att site; integrase; recombinase; ACes; gene therapy; transgenic animal;
KW      platform artificial chromosome expression system.
XX      OS      Diacrosoma sp.
OS      Synthetic.
XX      PN      WO200297059-A2.
XX      PD      05-DEC-2002.
PF      30-MAY-2002; 2002WO-US017452.
XX      PR      30-MAY-2001; 2001US-0294758P.
PR      21-MAR-2002; 2002US-0366891P.
XX      PA      (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.
XX      PI      Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;
PI      Stewart S, Shellard J;
XX      WP1, 2003-140461/13.
DR      N-PADB; ACC44649.
XX      Novel eukaryotic chromosome comprising one or many att sites which
PT      permits site-directed integration in the presence of lambda-integrase,
PT      useful for site-specific recombination-directed integration of DNA of
PT      interest.
XX      Example 1; Page 192-193; 272pp; English.
XX      The present invention describes a eukaryotic chromosome (I) comprising
CC      one or several att sites, where an att site is heterologous to the
CC      chromosome, and permits site-directed integration in the presence of
CC      lambda-integrase. Also described: (I) a platform artificial chromosome
CC      expression system (ACes) (II) comprising several sites that participate
CC      in recombinase catalysed recombination; and (2) a method (M1) for
CC      introducing a heterologous nucleic acid into a platform artificial
CC      chromosome. (I) can be used in gene therapy. (M1) is useful for
CC      introducing a heterologous nucleic acid molecule into a platform
CC      artificial chromosome, preferably an ACes. (II) is useful for producing a
CC      transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or
CC      mammal) by introducing (II) by cell fusion, lipid-mediated transfection
CC      by a carrier system, microinjection, microcell fusion, electroporation,
CC      microprojectile bombardment or direct DNA transfer into an embryonic
CC      cell, preferably a stem cell or an embryo. (II) comprises a heterologous
CC      nucleic acid that encodes a therapeutic product which is useful for
CC      making a library of ACes comprising random portions of a genome. ACC44612
CC      to ACC44732 and ABP96650 to ABP96657 represent sequences used in the
CC      exemplification of the present invention
XX      SQ      Sequence 230 AA;

Query Match          89.4%; Score 1085.5; DB 6; Length 230;
Best Local Similarity 88.1%; Pred. No. 3.1e-113;
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1

```

ID	Accession	Standard	Protein	Residue			
Db	1	MSCSKNV	IKFEMRFKVRMEGTVNGHEFEIKGEGRPYEGHCSVXKLMVTXGGLPFAFDI	60			
Qy	61	LSPQFGYSKYVVKHPADIPDYKKLSPEEGFKMERVMNPFEDGGVVTVTQDSLQDGCPIY		120			
Db	61	LSPQFGYSKYVVKHPADIPDYKKLSPEEGFKMERVMNPFEDGGVVTVTQDSLQDGCPIY		120			
Qy	121	KYKFIGVNPSPSDGPMQCKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI		180			
Db	121	KYKFIGVNPSPSDGPMQCKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI		180			
Qy	181	YMAKKP-VQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL		225			
Db	181	YMAKKP-VQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL		226			
RESULT 102							
ID	AAE00375	standard	protein	205 AA.			
AC	AAE00375						
DT	11-SEP-2003	(revised)					
DT	19-JUN-2001	(first entry)					
DE	Discosoma sp. red chromo/fluorescent protein, drFP583 (NFP-6).						
XX							
KW	Anthozoa; Chromoprotein; fluorescent protein; drFP583; NFP-6; sunscreen;						
KW	analyte detection assay; selectable marker; recombinant DNA application;						
KW	biosensor; pH indicator; invivo marker; selective filter.						
XX							
OS	Discosoma sp; red.						
XX							
FT	Key	Location/Qualifiers					
FT	Misc-difference	150..151					
FT	/note="Encoded by TTGTATCCTCGTGATGCGGTGTTGAAGAAGAGATT						
FT	CATAGGCTCTGAAGCTGAAGAAGCGTGGT"						
XX							
PN	WO200127150-A2.						
XX							
PD	19-APR-2001.						
XX							
PF	13-OCT-2000; 2000MO-US028477.						
XX							
PR	14-OCT-1999;	99US-00418529.					
PR	15-OCT-1999;	99US-00418917.					
PR	15-OCT-1999;	99US-00418922.					
PR	19-NOV-1999;	99US-00444338.					
PR	19-NOV-1999;	99US-00444341.					
PR	09-DEC-1999;	99US-00457556.					
PR	09-DEC-1999;	99US-00457898.					
PR	09-DEC-1999;	99US-00458144.					
PR	09-DEC-1999;	99US-00458477.					
PR	10-DEC-1999;	99MO-US029405.					
PR	14-JUN-2000;	2000US-0211607P.					
PR	14-JUN-2000;	2000US-0211609P.					
PR	14-JUN-2000;	2000US-0211626P.					
PR	14-JUN-2000;	2000US-0211627P.					
PR	14-JUN-2000;	2000US-0211687P.					
PR	14-JUN-2000;	2000US-0211766P.					
PR	14-JUN-2000;	2000US-0211880P.					
PR	14-JUN-2000;	2000US-0211888P.					
PR	14-JUN-2000;	2000US-0212070P.					
XX							
PA	(CLON-) CLONTECH LAB INC.						
XX							
PI	Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Tersikh A;						
XX							
DR	WPI; 2001-266409/27.						
XX							
DR	N-PSDB; AAD03615.						
XX							
PT	An Anthozoa chromo- or fluorescent protein (P1) present in an environment						
PT	other than its natural environment, useful as a label in analyte						







XX	DT	19-JUN-2001	(first entry)
XX	DE	Discosoma sp. red drFP583 (NFP-6) mutant E8 (N42H).	
XX	KW	Anthozoa; Chromoprotein; fluorescent protein; drFP583; NFP-6; sunscreen; analyte detection assay; selectable marker; recombinant DNA application; biosensor; pH indicator; invivo marker; selective filter; mutant; mutcin.	
XX	OS	Discosoma sp; red.	
XX	OS	Synthetic.	
XX	FT	Key	Location/Qualifiers
XX	FT	Misc-difference 42	/note= "wild type Asn substituted with His"
XX	PN	WO200127150-A2.	
XX	PD	19-APR-2001.	
XX	XX	13-OCT-2000; 2000WO-US028477.	
XX	PR	14-OCT-1999; 99US-00418529.	
XX	PR	15-OCT-1999; 99US-00418917.	
XX	PR	15-OCT-1999; 99US-00418922.	
XX	PR	19-NOV-1999; 99US-00444338.	
XX	PR	19-NOV-1999; 99US-00444341.	
XX	PR	09-DEC-1999; 99US-00457556.	
XX	PR	09-DEC-1999; 99US-00457898.	
XX	PR	09-DEC-1999; 99US-00458144.	
XX	PR	09-DEC-1999; 99US-00458477.	
XX	PR	10-DEC-1999; 99WO-US029405.	
XX	PR	14-JUN-2000; 2000US-0211607P.	
XX	PR	14-JUN-2000; 2000US-0211609P.	
XX	PR	14-JUN-2000; 2000US-0211626P.	
XX	PR	14-JUN-2000; 2000US-0211627P.	
XX	PR	14-JUN-2000; 2000US-0211687P.	
XX	PR	14-JUN-2000; 2000US-0211766P.	
XX	PR	14-JUN-2000; 2000US-0211880P.	
XX	PR	14-JUN-2000; 2000US-0211888P.	
XX	PR	14-JUN-2000; 2000US-0212070P.	
XX	PA	(CLON-) CLONTECH LAB INC.	
XX	PI	Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Tersikh A;	
XX	DR	WPI; 2001-266409/27.	
XX	PT	An Anthozoa chromo- or fluorescent protein (P1) present in an environment other than its natural environment, useful as a label in analyte detection assays and as a selectable marker in recombinant DNA applications.	
XX	PT	disclosure; Page; 69pp; English.	
XX	PS	The present sequence is a Discosoma sp. red chromo/fluorescent protein, drFP583 (NFP-6) mutant E8 (N42H). NFP-6 is present in an environment other than its natural environment and has an absorbance maximum ranging from 250nm to 750nm and more usually from 540nm to 580nm and emission maximum ranging from 275nm to 775nm and more usually from 565 to 605nm.	
XX	CC	The chromoproteins or fluorescent proteins are useful as labels in analyte detection assays, as selectable markers in recombinant DNA applications, as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator and as in vivo markers in animals. They are also useful in sunscreens and as selective filters. Note: The present sequence is not shown in the specification, but derived from drFP583 (NFP-6) sequence (AAE00375) shown in figure 6	
XX	CC	(AAE00375) shown in figure 6	
XX	CC	Sequence 205 AA;	
XX	XX	Query Match 86.7%; Score 1053; DB 4; Length 205;	
XX	XX	Best Local Similarity 88.0%; Pred. No. 1.2e-109;	
XX	XX	Matches 198; Conservative 3; Mismatches 4; Indels 20; Gaps 1;	

```

QY      1 MRSSKKNVKEFMRFFKVRMEGTVNGHFEFLGEGEGRPYEGHNTVKLKVTKGADLPFAWDI 60
Db      1 MRSSKKNVKEFMRFFKVRMEGTVNGHFEFLGEGEGRPYEGHNTVKLKVTKGADLPFAWDI 60
QY      61 LSPQFGYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFGYGSKYVYKH PADIPDYKCLSPFEGFKMERVMNFEDGGVVTVTQDSSLQDGCFTY 120
QY      121 KVKFIGNVFPDGPVVMQKTMGWEASTERLYPRDGLKGBIHKALKKDGCHYLVEFKSI 180
Db      121 KSSSLALNFPDGPVVMQKTMGWEASTERL-----GCHYLVEFKSI 160
QY      181 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYEQYERTEGRHHLFL 225
Db      161 YMAKKPVQLPGYVVVDSKLDITSHNEDYTYEQYERTEGRHHLFL 205

RESULT 105
ADC24133
ID      ADC24133 standard; protein; 225 AA.
XX
AC      ADC24133;
XX
DT      18-DEC-2003 (first entry)
XX
DE      Discosoma red fluorescent protein variant mRFP1.
XX
KW      Discosoma red fluorescent protein; DsRed; AB interface; AC interface;
KW      fluorescent protein variant; transcription induction detection;
KW      fluorescence energy resonance transfer; FRET; protein kinase;
KW      protein phosphatase; ion indicator; mutant; mRFP1; muteln.
XX
OS      Synthetic.
OS      Discosoma.
XX
PN      US2003059835-A1.
XX
PD      27-MAR-2003.
XX
PF      10-APR-2002; 2002US-00121258.
XX
PR      26-FEB-2001; 2001US-00794308.
PR      24-MAY-2001; 2001US-00866538.
XX
PA      (TSIE/) TSIE N Y.
PA      (CAMP/) CAMPBELL R E.
XX
PI      Tsien RY, Campbell RE;
XX
DR      WPI; 2003-743764/70.
DR      N-PSDB; ADC24132.
XX
PT      Novel polynucleotide sequence encoding Discosoma red fluorescent protein
PT      variant having a reduced propensity to oligomerize, useful for detecting
PT      transcriptional activity.
XX
PS      Disclosure; SEQ ID NO 8; 67pp; English.
XX
CC      The invention describes a polynucleotide sequence (I) encoding a
CC      Discosoma red fluorescent protein (DsRed) variant having a reduced
CC      propensity to oligomerize, comprising amino acid substitutions at the AB
CC      and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225
CC      amino acids, given in the specification, where the substitutions result
CC      in reduced propensity of the DsRed variant to form tetramers. (I) is
CC      useful for detecting transcriptional activity by providing a host cells
CC      conatinng a vector which comprises (I) operatively linked to an
CC      expression control sequence, and an unit to assay the variant fluorescent
CC      protein fluorescence, and assaying fluorescence of the variant
CC      fluorescent protein produced by (VII), where variant fluorescent protein
CC      fluorescence is indicative of transcriptional activity. A polynucleotide
CC      encoding a fusion protein is useful for the analysis of in vivo
CC      localisation or trafficking of a polypeptide of interest. A polypeptide

```







```
PA      (MACD/) MACDONALD M L.  
PA      (LAME/) LAMERDIN J.  
XX  
PI      Watson Michnick SW, Macdonald ML, Lamerdin J;  
XX  
DR      WPI; 2004-533363/51.  
DR      N-PSDB; ADQ59565.  
PT  
PT      Composition useful in protein fragment complementation assays for drug  
PT      discovery and high-throughput screening, comprising complementary  
PT      fragments of protein or mutant protein, generating optically detectable  
PT      signal when associated.  
XX  
PS      Disclosure; SEQ ID NO 16; 34pp, English.  
XX  
CC      The invention describes a composition (I) comprising complementary  
CC      fragments of a protein or mutant protein, where the fragments generate an  
CC      optically detectable signal when associated, and each of the mutant  
CC      protein fragments is fused to a separate molecule. Also described are:  
CC      protein fragment complementation assays (PCAs) for detection of molecular  
CC      interactions, involving reassembling separate fragments from an optically  
CC      detectable protein, and detecting the reassembly by units of  
CC      reconstitution of activity of the optically detectable protein, where the  
CC      reassembly of the fragments is operated by the interaction of molecular  
CC      domains fused to each fragment, and is independent of other molecular  
CC      processes; detecting biomolecular interaction, involving selecting an  
CC      appropriate optically detectable protein, effecting fragmentation of the  
CC      optically detectable protein such that the fragmentation results in  
CC      reversible loss of protein function, fusing or attaching fragments of the  
CC      optically detectable protein separately to other molecules, reassociating  
CC      the protein fragments through interactions of the molecules that are  
CC      fused or attached to the fragments, and detecting the resulting optical  
CC      signal; designing and engineering of PCAs based on fluorescent protein;  
CC      and a method and composition for the construction of multi-color PCAs.  
CC      (I) is useful in PCAs and other assays for drug discovery, target  
CC      validation, high-throughput screening, high-content screening, pathway  
CC      mapping, drug mechanism-of-action studies, biosensors and diagnostics.  
CC      (I) is useful for engineering different colour PCAs for a variety of  
CC      applications in biology and biotechnology. This is the amino acid  
CC      sequence of a Discosoma monomerised red fluorescent protein from which  
CC      PCA fragments can be isolated.  
XX  
SQ      Sequence 225 AA;  
  
Query Match          84.1%; Score 1021; DB 8; Length 225;  
Best Local Similarity 86.9%; Pred. No. 5.6e-106;  
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0  
  
QY      1 MRSSKNVIEKEMRFKVRMEGTVGNGHEFELEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
    ||::|||  
Db       1 MASSEDVIKEFMRFKVRMEGSVNGHEFELEGGEGRPRYEGTQTAKLVTKGGLPPFAMD I 60  
  
QY      61 LSPOFGYSKVYVKHPADIPDYKKLSPEEFGKMERVMNPFEDSGVVTTODSSLQDGCFIY 120  
    |||  
Db       61 LSPFOFGYSKAYVVKHPADIPDYKLSPFEFGKMERVMNPFEDSGVVTTQDSSLQDGEFIY 120  
  
QY      121 KYKFIGVNPPSDGPVMOKKTMGWEASTERLYPRDVLKGEIHKALKKDGHYLVEFXSI 180  
    |||  
Db       121 KYKLRGTNFPSDGPVMOKKTMGWEASTERYPEDGALKGEIKMKRLKLDGGHYDAEVKTT 180  
  
QY      181 YMAKKPVQLPGYYYVDSKLDITSHNEDYTIVEQYERTEGRH 221  
    |||  
Db       181 YMAKKPVQLPGAYKTDIKLDITSHNEDYTIVEQYERAEGRH 221  
  
RESULT 108  
AEA54910 standard; protein; 747 AA.  
XX  
AC      AEA54910;  
XX  
DT      25-AUG-2005 (first entry)  
XX
```

```
DE Fluorescence-related EGFP-caspase-3 DEVD linker-mRFP(x2) fusion protein.
XX fluorescence; protein localization; cellular transport; organelle;
KW mitochondria; fusion protein; enhanced green fluorescent protein; EGFP;
KM caspase-3.
XX Unidentified.
OS Chimeric.
PN WO2005054464-A1.
XX
PD 16-JUN-2005.
XX
PF 03-DEC-2004; 2004WO-JP018437.
XX
PR 03-DEC-2003; 2003JP-00404472.
PR 27-JAN-2004; 2004JP-00018344.
XX
PA (RIKE ) RIKEN KK.
PA (MED1-) MEDICAL & BIOLOGICAL LAB CO LTD.
PI Miyawaki A, Kogure T, Hama H, Kinjo M, Saito K, Karasawa S;
PI Araki T;
XX WPI; 2005-444979/45.
XX N-PSDB; AEA54911.
DR
PS Example 10; SEQ ID NO 49; 218pp; Japanese.
XX
CC The invention relates to a novel fluorescent protein existing as a
CC monomer and comprising any one of SEQ ID No. 1-29 (odd SEQ ID numbers)
CC given in the specification, or any one of the sequences chosen from SEQ
CC ID No. 1-29 (odd SEQ ID numbers), with one or more amino acid
CC substitutions, deletions and/or additions, and having fluorescent
CC characteristics. The fusion protein of the invention may be useful for
CC analyzing the location or movement of a protein in a cell in vivo and for
CC labeling and analyzing the location or movement of intracellular
CC organelles, such as mitochondria. The current sequence is that of the
CC fluorescence-related enhanced green fluorescent protein (EGFP)-caspase-3
CC DEVD linker-mRFP(x2) fusion protein of the invention.
XX
SQ Sequence 747 AA;

Query Match      84.1%; Score 1021; DB 9; Length 747;
Best Local Similarity 86.9%; Pred. No. 3.3e-105;
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0

QY   1 MRSSKNVKEFMRFKVMEGTVNGHPEIEGEGGRPYEGHNTVKLKYTKGGPLPFAMDI 60
    |||::|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:
Db   269 MASSEDVIKEFMRFKVMESGVNGHBEIIEGEGGRPYEGTAKLVKTKGGPLPFAMDI 328

QY   61 LSPFOFGYSKVYVKHPADIPDYKKLSFPEGFKWERVMNFBDDGVVTVTQDSSLQDGCFFIY 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   329 LSPFOFGYSKAYVKHPADIPDYKLKLSPEGFKWERVMNFBDDGVVTVTQDSSLQDGEFIY 388

QY   121 KVKFIVGNFSPSDGPVWQKTGMWEASTERLYPRDGLKGELHKALKLDGGHYLVVEPKSI 180
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | :
Db   389 KVKLRGTNFPSDGPVWQKTGMWEASTERMYPEDGALKEGIKRLKLDGGHYDAEVKTT 448

QY   181 YMAKKPVQLPGYYVVDSKLDTSHNEDYTIVEQYERTEGRH 221
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db   449 YMAKKPVQLPGAYKTDIKLDTSHNEDYTIVEQYERAEGRH 489

RESULT 109
ADL46281
ID ADL46281 standard; protein; 225 AA.
XX
AC ADL46281;
```



XX 20-MAY-2004 (first entry)  
DT  
XX  
DE Discosoma red fluorescent protein (DsRed) variant mRFP1.1 coding sequence.  
XX  
KW red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.  
OS Synthetic.  
PN WC2003086446-A1.  
XX  
PD 23-OCT-2003.  
XX  
PD 09-APR-2003; 2003WO-US010879.  
XX  
XX 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
XX Tsien RY, Campbell RE, Baird GS;  
PI WPI, 2003-845265/78.  
DR  
XX  
XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
XX Claim 70; SEQ ID NO 79; 166pp; English.  
PS  
XX  
XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC variant mRFP1.1 protein which contains a Q66M and a T147S mutation.  
XX  
XX Sequence 225 AA;  
XX

[illegible]

Search completed: January 11, 2006, 02:08:28  
Job time : 81 secs



**THIS PAGE BLANK (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 11, 2006, 02:05:28 ; Search time 16 Seconds  
(without alignments)  
1353.048 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : PIR 80:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
No matches found					

No matches found

Search completed: January 11, 2006, 02:10:16  
Job time : 16 secs



**THIS PAGE BLANK (USPTO)**



Gencore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:04:58 ; Search time .73 Seconds  
(without alignments)  
2174.574 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Uniprot 05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	1 RFP DISSP	Q9U6Y8 discosoma s
2	1196	98.5	236	2 Q66ND7_9CNID	Q66nd7 discosoma s
3	1193	98.3	225	2 Q6KF85_9CNID	Q6kf85 discosoma s
4	1191	98.1	236	2 Q66ND8_9CNID	Q66nd8 discosoma s
5	1085.5	89.4	230	2 Q9GTJ7_9CNID	Q9gtj7 discosoma s
6	985	81.1	226	2 Q5S3G8_9CNID	Q5s3g8 discosoma s
7	972	80.1	226	2 Q5S3G7_9CNID	Q5s3g7 discosoma s

ALIGNMENTS

RESULT 1

RFP DISSP STANDARD; PRT; 225 AA.

ID RFP DISSP

AC Q9U6Y8;

DT 01-FEB-2005 (Rel. 46, Created)

DT 01-FEB-2005 (Rel. 46, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Red fluorescent protein dFpP583 (DsRed).

OS Discosoma sp. (Sea anemone).

OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.

OC NCBI\_TaxID=86600;

OX [1]

RN NUCLEOTIDE SEQUENCE.

RP MEDLINE=99436614; PubMed=10504696; DOI=10.1038/13657;

RX Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,

RA Markelov M.L., Lukyanov S.A.;

RT "Fluorescent proteins from nonbioluminescent Anthozoa species.;"

RL Nat. Biotechnol. 17:969-973(1999).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) .

RX PubMed=11101896; DOI=10.1038/81992;

RA Wall M.A., Socolich M., Ranganathan R.;

RT "The structural basis for red fluorescence in the tetrameric GFP

RT homolog DsRed.;"

RL Nat. Struct. Biol. 7:1133-1138(2000).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND SUBUNIT.

RX PubMed=11209050; DOI=10.1073/pnas.98.2.462;

RA Yarbrough D., Wachter R.M., Kallio K., Matz M.V., Remington S.J.;

RT "Refined crystal structure of DsRed, a red fluorescent protein from

RT coral, at 2.0-A resolution.;"

RT Proc. Natl. Acad. Sci. U.S.A. 98:462-467(2001).

CC -1- FUNCTION: Thought to play a role in photoprotection of the coral's

CC resident symbiont microalgae's photosystems from photoinhibition

CC caused by high light levels found near the surface of coral reefs.

CC In deeper water, the fluorescence may be to convert blue light

CC into longer wavelengths more suitable for use in photosynthesis by

CC the microalgal symbionts.

CC -1- BIOPHYSICOCHEMICAL PROPERTIES:

CC Absorption:

CC Abs(max)=558 nm;

CC Note=Exhibits a smaller absorbance peak at 494 nm. The broad

CC fluorescence emission spectrum peaks at 583 nm;

CC -1- SUBUNIT: Homotetramer.

CC -1- PTM: Contains a chromophore consisting of modified amino acid

CC residues. The chromophore is formed by autocatalytic backbone

CC condensation between Xaa-N and Gly-(N+2), oxidation of Tyr-(N+1)

CC to dihydroxyrosine, and formation of a double bond to the alpha-

CC amino nitrogen of residue Xaa-N. Maturation of the chromophore

CC requires nothing other than molecular oxygen.

CC -1- BIOTECHNOLOGY: Fluorescent proteins have become a useful and

CC ubiquitous tool for making chimeric proteins, where they function

CC as a fluorescent protein tag. Typically they tolerate N- and C-

CC terminal fusion to a broad variety of proteins. They have been

CC expressed in most known cell types and are used as a noninvasive

CC fluorescent marker in living cells and organisms. They enable a

CC wide range of applications where they have functioned as a cell

CC lineage tracer, reporter of gene expression, or as a measure of

CC protein-protein interactions.

CC -1- SIMILARITY: Belongs to the GFP family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

CC EMBL; AF168419; AAF03369.1; -; mRNA.

CC PDB; 1G7K; X-ray; A/B/C/D=2-225.

CC PDB; 1G8X; X-ray; A/B/C/D=1-225.

CC InterPro; IPR011584; GFP\_related.

CC InterPro; IPR000786; Green\_fl\_protein.

CC Pfam; PF01353; GFP; 1.

CC PRINTS; PR01229; GFP; 1.

CC PRODOM; PD013756; Green\_fl\_protein; 1.

CC 3D-structure; Chromophore; Luminescence; Photoprotein.

CC MOD RES 67 67 (Z)-2,3-dihydroxyrosine.

CC CROSSLINK 66 68 2-iminomechyl-5-imidazolinone (Gln-Gly).

CC SEQUENCE 225 AA; 25931 MW; FB9A5369778F689 CRC64;

CC -----

Query Match 100.0%; Score 1214; DB 1; Length 225;

Best Local Similarity 100.0%; Pred. No. 1e-96;

Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTYKGGLPFAWDI 60

Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTYKGGLPFAWDI 60

QY 61 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWRVMNPFEDGVVTVTQDSIQDGCFTY 120

Db 61 LSPQFOYGSKVYVGHADIPDYKLSFPEGFKWRVMNPFEDGVVTVTQDSIQDGCFTY 120



```

QY      121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKGEIHKALKLKDGGHYLVEFKSI 180
      |||
Db      121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKGEIHKALKLKDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
      |||
Db      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 2
Q66ND7_9CNID PRELIMINARY; PRT; 236 AA.
ID Q66ND7_9CNID PRELIMINARY; PRT; 236 AA.
AC Q66ND7;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Enhanced red fluorescent protein R+.
OS Discosoma sp. RC-2004.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=289055;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT "Cloning of Anthozoan Fluorescent Protein Genes.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY679107; AAU04444.1; -; mRNA.
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PRO1229; GFLUORESCENT.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 236 AA; 27032 MW; BB3844BE06829EF0 CRC64;

Query Match 98.5%; Score 1196; DB 2; Length 236;
Best Local Similarity 98.2%; Pred. No. 3.9e-95;
Matches 221; Conservative 2; Mismatches 2; Indels 0; Gaps 0

QY      1 MRSSKNVKEFMRFKVRMEGTIVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60
      |||
Db      1 MSCSKNVKEFMRFKVRMEGTIVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY      61 LSPQFOYGSKVYVKHPADIPDYKLSPEEGFKMERVMNFEDGVVTVTQDSLQDGCFTY 120
      |||
Db      61 LSPQFOYGSKVYVKHPADIPDYKLSPEEGFKMERVMNFEDGVVTVTQDSLQDGCFTY 120

QY      121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKGEIHKALKLKDGGHYLVEFKSI 180
      |||
Db      121 KVKFIGVNFPSDGPVMQKKTGMWEASTERLYPRDGLVKGEIHKALKLKDGGHYLVEFKTI 180

QY      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
      |||
Db      181 YMAKKPVQLPGYIYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3
Q6KF85_9CNID PRELIMINARY; PRT; 225 AA.
ID Q6KF85_9CNID PRELIMINARY; PRT; 225 AA.
AC Q6KF85;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Orange fluorescent protein FP586.
OS Discosoma sp. JW-2002.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=208461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wiedenmann J., Girod A.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF545828; AAQ11987.1; -; mRNA.
DR HSSP; P42212; 1B9C.
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP_UORESCENT.
DR Prodom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 225 AA; 25791 MW; E151D0E497AA23FA CRC64;

Query Match 98.3%; Score 1193; DB 2; Length 225;
Best Local Similarity 98.2%; Pred. No. 6.7e-95;
Matches 221; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLTGKGPLPFAWDI 60
DB 1 MSCSKNVIKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLTGKGPLPFAWDI 60
QY 61 LSPFOYGSKVYVKHPADIPDYKCLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120
DB 61 LSPFOYGSKVYVKHPADIPDYKCLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120
QY 121 KYKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLVKGELIKALKLDGGHYLVBFKSI 180
DB 121 KYKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLVKGELIKALKLDGGHYLVBFKSI 180
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225
DB 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTYVEQYERTEGRHHLFL 225

RESULT 4
Q66ND8_9CNID PRELIMINARY; PRT; 236 AA.
ID Q66ND8; AC
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Red fluorescent protein R1.
OS Discosoma sp. RC-2004.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;
OC Discosomatidae; Discosoma.
OX NCBI_TaxID=289055;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Carter R.W., Gibbs P.D.L., Schmale M.C.;
RT "Cloning of Anthozoan Fluorescent Protein Genes.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY679106; AAU04443.1; -; mRNA.
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP_UORESCENT.
DR Prodom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 236 AA; 27042 MW; BB3844BCD6834EF3 CRC64;

Query Match 98.1%; Score 1191; DB 2; Length 236;
Best Local Similarity 97.8%; Pred. No. 1.1e-94;
Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLTGKGPLPFAWDI 60
DB 1 MSCSKNVIKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLTGKGPLPFAWDI 60
QY 61 LSPFOYGSKVYVKHPADIPDYKCLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120
DB 61 LSPFOYGSKVYVKHPADIPDYKCLSFPEGFKMERVMNFEDGGVVTYTQDSSLQDGCFTY 120
QY 121 KYKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLVKGELIKALKLDGGHYLVBFKSI 180
DB 121 KYKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLVKGELIKALKLDGGHYLVBFKSI 180

```



Qy 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 5

Q9GTJ7\_9CNID PRELIMINARY; PRT; 230 AA.  
AC Q9GTJ7;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2003 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Red fluorescent protein.  
GN Name=FP593;  
OS Discosoma sp. SSAL-2000.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=137428;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20434599; PubMed=10981720; DOI=10.1016/S0014-5793(00)01895-0;  
RA Fradkov A.P., Chen Y., Ding L., Barbova E.V., Matcz M.V.,  
RA Lukyanov S.A.;  
RT "Novel fluorescent protein from Discosoma coral and its mutants  
possesses a unique far-red fluorescence.";  
RL FEBS Lett. 479:127-130(2000).  
DR EMBL; AF272711; AAC16224.1; -; mRNA.  
DR HSSP; Q9U6Y8; 1GGX.  
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
DR InterPro; IPR011584; GFP related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR SEQUENCE 230 AA; 26370 MW; 5215B1B436D67E51 CRC64;

Query Match 89.4%; Score 1085.5; DB 2; Length 230;  
Best Local Similarity 88.1%; Pred. No. 1.3e-85;  
Matches 199; Conservative 17; Mismatches 9; Indels 1; Gaps 1;

Qy 1 MRSSKVNIFEKMRFRKVRMEGTVNGHFEIEGEGRPYEGHNTVTKLVTKGGLPFAWDI 60  
Db 1 MSCSKNVIFEKMRFRKVRMEGTVNGHFEIEKGEGRPYEGHCSVKLMVTKGGLPFAFDI 60  
Qy 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLKDGCFY 120  
Qy 121 KVKPIGVNFPSPDGPVQKKTGMWEASTERLYPRDGLKGEIHKALKDKDGHYLVFEKSI 180  
Db 121 EVKPIGVNFPSPDGPVQKRTRGWEASSERLYPRDGLKGDIMALRLGEGGHYLVFEKSI 180  
Qy 181 YMAKKP-VQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKP-VQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHPFI 226

RESULT 6

Q5S3G8\_9CNID PRELIMINARY; PRT; 226 AA.  
AC Q5S3G8;  
DT 01-FEB-2005 (Tremblrel. 29, Created)  
DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)  
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)  
DE Fluorescent protein raspberry.  
OS Discosoma sp. LW-2004.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=301246;

Qy 11  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15556995; DOI=10.1073/pnas.0407752101;  
RA Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y.;  
RT "Evolution of new nonantibody proteins via iterative somatic

RT hypermutation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
DR EMBL; AY786536; AAV65486.1; -; mRNA.  
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
DR SEQUENCE 226 AA; 25513 MW; 159A8FBE2D453A5B CRC64;

Query Match 81.1%; Score 985; DB 2; Length 226;  
Best Local Similarity 85.3%; Pred. No. 6.3e-77;  
Matches 185; Conservative 6; Mismatches 26; Indels 0; Gaps 0;

Qy 5 KNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVTKLVTKGGLPFAWDILSPQ 64  
Db 6 EEVKEFMRFKVRMEGSVNGHFEIEGEGGRPYEGTQTALKVTKGGLPFAWDILSPQ 65  
Qy 65 FQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLQDGCFTYKVF 124  
Db 66 CMYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLQDGEFTYKVKL 125  
Qy 125 IGVNFPSPDGPVQKKTGMWEASTERLYPRDGLKGEIHKALKDKDGHYLVFEKSIYMAK 184  
Db 126 RGTNFPSPDGPVQKKTGMWEASSERNYPEDGALKGEMKRLKDKDGHYDARVKTYYMAK 185  
Qy 185 KPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 186 KPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 222

RESULT 7

Q5S3G7\_9CNID PRELIMINARY; PRT; 226 AA.  
AC Q5S3G7;  
DT 01-FEB-2005 (Tremblrel. 29, Created)  
DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)  
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)  
DE Fluorescent protein plum.  
OS Discosoma sp. LW-2004.  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia;  
OC Discosomatidae; Discosoma.  
OX NCBI\_TaxID=301246;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15556995; DOI=10.1073/pnas.0407752101;  
RA Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y.;  
RT "Evolution of new nonantibody proteins via iterative somatic  
hypermutation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:16745-16749(2004).  
DR EMBL; AY786537; AAV65487.1; -; mRNA.  
DR SMR; Q5S3G7; 8-222.  
DR GO; GO:0006091; P:generation of precursor metabolites and energy; IEA.  
DR InterPro; IPR011584; GFP\_related.  
DR InterPro; IPR000786; Green\_fl\_protein.  
DR Pfam; PF01353; GFP; 1.  
DR PRINTS; PR01229; GFLUORESCENT.  
DR ProDom; PD013756; Green\_fl\_protein; 1.  
DR SEQUENCE 226 AA; 25530 MW; 6DF38CAD2AB28BED CRC64;

Query Match 80.1%; Score 972; DB 2; Length 226;  
Best Local Similarity 83.4%; Pred. No. 8.4e-76;  
Matches 181; Conservative 8; Mismatches 28; Indels 0; Gaps 0;

Qy 5 KNVKEFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVTKLVTKGGLPFAWDILSPQ 64  
Db 6 EEVKEFMRFKHEMGSVNGHFEIEGEGGRPYEGTQTARLKVTKGGPLPFAWDILSPQ 65  
Qy 65 FQYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLQDGCFTYKVF 124  
Db 66 IMYGSKVYVKHPADIPDYKLSFPEGFKWERVMNFEDEGGVVTVTQDSSLQDGEFTYKVKV 125



Qy	125	IGVNFPSDGPVMOKKTMGWEASTERLYPRDGLKGEIHKALKDKGHHYLVFESITYMAK	184
Db	126	RGTNFPSPDGPVMOKKTMGWEASSERMYPEDGALKGEMKMRLRLKDGHHYDAEVKITYMAK	185
Qy	185	KPVOLPGYYVDSKLDITSHNEDYTYVEQYERTGRH	221
Db	186	KPVOLPGAYKTDIKLDITSHNEDYTYVEQYERABGRH	222

Search completed: January 11, 2006, 02:09:49  
Job time : 74 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:06:44 ; Search time 23 Seconds  
(without alignments)  
808.784 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/H\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/RE\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	2	US-09-866-538-12
2	1214	100.0	225	2	US-09-865-291-12
3	1210	99.7	240	2	US-10-152-296-2

ALIGNMENTS

RESULT 1  
US-09-866-538-12  
; Sequence 12, Application US/09866538  
; Patent No. 6852849  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/09/866,538  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-866-538-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFGYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGWGEASTERLYPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKKTGWGEASTERLYPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTGRHHLFL 225  
DB 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTGRHHLFL 225

RESULT 2

US-09-865-291-12  
; Sequence 12, Application US/09865291  
; Patent No. 6900304  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: TING, Alice  
; APPLICANT: ZHANG, Jin  
; TITLE OF INVENTION: EMISSION RATIONETRIC INDICATORS OF PHOSPHORYLATION  
; FILE REFERENCE: REGEN1550  
; CURRENT APPLICATION NUMBER: US/09/865,291  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 12  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
US-09-865-291-12

Query Match 100.0%; Score 1214; DB 2; Length 225;  
Best Local Similarity 100.0%; Pred. No. 3e-132;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
DB 1 MRSSKNVKEFMRFKVRMEGTNGHEFEIEGEGEGRPYEGHNTVKLKVTGKGPLPFAWDI 60  
QY 61 LSPQFGYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYGSKYVVKHPADIPDYKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KYKFIGVNFPSDGPVMQKKTGWGEASTERLYPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
DB 121 KYKFIGVNFPSDGPVMQKKTGWGEASTERLYPRDGLKGEIHKALKKDGGHYLVFCKSI 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTGRHHLFL 225  
DB 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIQEYERTGRHHLFL 225

RESULT 3

US-10-152-296-2  
; Sequence 2, Application US/10152296  
; Patent No. 6723537  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US



; CURRENT APPLICATION NUMBER: US/10/152,296
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/291,871
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mammalian
; OTHER INFORMATION: codon-optimized variant (DSRED) of Discosoma sp.
; OTHER INFORMATION: "red" red fluorescent protein (RFP)
US-10-152-296-2

Query Match 99.7%; Score 1210; DB 2; Length 240;
Best Local Similarity 99.6%; Pred. No. 9.5e-132;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 60
:|||||
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMD 61
61 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120
|||||
Db 62 LSPQFGYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 121
121 KVKFPGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180
|||||
Db 122 KVKFPGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 181
181 YMAKKEVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 225
|||||
Db 182 YMAKKEVQLPGYYVDSKLDITSHNEDYTIVEQYERTEGRHHLFL 226

Search completed: January 11, 2006, 02:10:50
Job time : 23 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:07:04 ; Search time 62 Seconds  
(without alignments)  
1516.316 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues  
Total number of hits satisfying chosen parameters: 75

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Published Applications AA Main:\*

- 1: /cgnt2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgnt2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 3: /cgnt2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 4: /cgnt2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 5: /cgnt2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 6: /cgnt2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	225	3	US-09-999-745-67 Sequence 67, Appl
2	1214	100.0	225	3	US-09-866-538-12 Sequence 12, Appl
3	1214	100.0	225	3	US-09-794-308-12 Sequence 12, Appl
4	1214	100.0	225	3	US-09-865-291-12 Sequence 12, Appl
5	1214	100.0	225	4	US-10-006-922-12 Sequence 12, Appl
6	1214	100.0	225	4	US-10-081-864-8 Sequence 8, Appl
7	1214	100.0	225	4	US-10-121-258-1 Sequence 1, Appl
8	1214	100.0	225	4	US-10-315-920-2 Sequence 2, Appl
9	1214	100.0	225	4	US-10-132-067-4 Sequence 4, Appl
10	1214	100.0	225	4	US-10-370-570-56 Sequence 56, Appl
11	1214	100.0	225	4	US-10-406-618-32 Sequence 32, Appl
12	1214	100.0	225	4	US-10-433-640-13 Sequence 13, Appl
13	1214	100.0	225	4	US-10-311-030-7 Sequence 7, Appl
14	1214	100.0	225	5	US-10-845-484-3 Sequence 3, Appl
15	1214	100.0	225	5	US-10-885-988-12 Sequence 12, Appl
16	1214	100.0	225	5	US-10-656-029-22 Sequence 22, Appl
17	1214	100.0	225	5	US-10-857-622-12 Sequence 12, Appl
18	1214	100.0	225	5	US-10-505-486-26 Sequence 26, Appl
19	1214	100.0	225	5	US-10-844-064A-2 Sequence 2, Appl
20	1214	100.0	225	5	US-10-931-304-1 Sequence 1, Appl
21	1214	100.0	487	4	US-10-343-977-1 Sequence 1, Appl
22	1214	100.0	487	5	US-10-343-977-1 Sequence 1, Appl
23	1214	100.0	506	4	US-10-343-977-2 Sequence 2, Appl
24	1214	100.0	506	5	US-10-343-977-2 Sequence 2, Appl
25	1214	100.0	547	4	US-10-343-977-3 Sequence 3, Appl
26	1214	100.0	547	5	US-10-343-977-3 Sequence 3, Appl
27	1210	99.7	226	4	US-10-311-030-9 Sequence 9, Appl

28	1210	99.7	240	4	US-10-152-296-2	Sequence 2, Appl
29	1210	99.7	240	4	US-10-739-656-2	Sequence 2, Appl
30	1210	99.7	545	4	US-10-214-932-52	Sequence 52, Appl
31	1210	99.7	548	4	US-10-214-932-76	Sequence 76, Appl
32	1207	99.4	225	4	US-10-121-258-20	Sequence 20, Appl
33	1207	99.4	225	4	US-10-315-920-4	Sequence 4, Appl
34	1207	99.4	225	5	US-10-931-304-20	Sequence 20, Appl
35	1196	98.5	236	4	US-10-314-936-4	Sequence 2, Appl
36	1196	98.5	236	6	US-10-314-936-4	Sequence 2, Appl
37	1196	98.5	236	6	US-11-021-014-2	Sequence 4, Appl
38	1196	98.5	236	6	US-11-021-014-4	Sequence 4, Appl
39	1195.5	98.5	226	4	US-10-724-178-12	Sequence 12, Appl
40	1191	98.1	225	4	US-10-315-920-6	Sequence 6, Appl
41	1186	97.7	225	4	US-10-442-148A-7	Sequence 7, Appl
42	1186	97.7	239	4	US-10-442-148A-8	Sequence 8, Appl
43	1160	95.6	225	4	US-10-121-258-4	Sequence 4, Appl
44	1160	95.6	225	5	US-10-931-304-4	Sequence 4, Appl
45	1153	95.0	225	4	US-10-121-258-24	Sequence 24, Appl
46	1153	95.0	225	5	US-10-931-304-24	Sequence 24, Appl
47	1132	93.2	225	4	US-10-423-688A-41	Sequence 41, Appl
48	1121	92.3	226	4	US-10-121-258-6	Sequence 6, Appl
49	1121	92.3	226	5	US-10-931-304-6	Sequence 6, Appl
50	1121	92.3	602	6	US-11-052-001-6	Sequence 6, Appl
51	1119	92.2	225	4	US-10-006-922-44	Sequence 44, Appl
52	1119	92.2	225	4	US-10-081-864-12	Sequence 12, Appl
53	1095	90.2	234	5	US-10-931-304-81	Sequence 81, Appl
54	1095	90.2	476	5	US-10-931-304-106	Sequence 106, Appl
55	1085.5	89.4	230	4	US-10-006-922-18	Sequence 18, Appl
56	1085.5	89.4	230	4	US-10-161-403-40	Sequence 40, Appl
57	1085.5	89.4	230	6	US-11-006-076-40	Sequence 40, Appl
58	1046	86.2	205	4	US-10-006-922-46	Sequence 46, Appl
59	1021	84.1	225	4	US-10-121-258-8	Sequence 8, Appl
60	1021	84.1	225	4	US-10-724-178-16	Sequence 16, Appl
61	1021	84.1	225	5	US-10-931-304-8	Sequence 8, Appl
62	1014	83.5	225	5	US-10-931-304-108	Sequence 108, Appl
63	1012	83.4	225	5	US-10-931-304-79	Sequence 79, Appl
64	1007	82.9	225	5	US-10-931-304-104	Sequence 104, Appl
65	1002	82.5	225	5	US-10-931-304-100	Sequence 100, Appl
66	994	81.9	236	5	US-10-931-304-85	Sequence 85, Appl
67	992	81.7	236	5	US-10-931-304-89	Sequence 89, Appl
68	989	81.5	236	5	US-10-931-304-102	Sequence 102, Appl
69	988	81.4	236	5	US-10-931-304-83	Sequence 83, Appl
70	987	81.3	236	5	US-10-931-304-92	Sequence 92, Appl
71	986	81.2	235	5	US-10-931-304-94	Sequence 94, Appl
72	973	80.1	236	5	US-10-931-304-96	Sequence 96, Appl
73	973	80.1	236	5	US-10-931-304-98	Sequence 98, Appl
74	960	79.1	236	5	US-10-931-304-87	Sequence 87, Appl
75	848	69.9	183	4	US-10-724-178-1041	Sequence 1041, Ap

ALIGNMENTS

RESULT 1  
US-09-999-745-67  
; Sequence 67, Application US/09999745  
; Patent No. US20020157120A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REAGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: REGENT1470-1  
; CURRENT APPLICATION NUMBER: US/09/999,745  
; CURRENT FILING DATE: 2001-10-23  
; PRIOR APPLICATION NUMBER: 09/316,920  
; PRIOR FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 67  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.







```
RESULT 5
US-10-006-922-12
; Sequence 12, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Pradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Teretskikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/fluorophores and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma species
US-10-006-922-12

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120

QY      121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180
        |||||||
DB      121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||||||
DB      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 6
US-10-081-864-8
; Sequence 8, Application US/10081864
; Publication No. US20030022287A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Yanushkevich, Yuriy
; APPLICANT: Savitsky, Alexandr
; APPLICANT: Pradkov, Arcady
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-067
; CURRENT APPLICATION NUMBER: US/10/081,864
; CURRENT FILING DATE: 2002-06-19
```

```
; PRIOR APPLICATION NUMBER: 10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/270,983
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-081-864-8

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120

QY      121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180
        |||||||
DB      121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
        |||||||
DB      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
```

```
RESULT 7
US-10-121-258-1
; Sequence 1, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(225)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-1

Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60
        |||||||
DB      1 MRSSKNVKEFMRFKVMNEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60

QY      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
        |||||||
DB      61 LSPQFOYGSKVYVKHPADIPDYKLSFPEGFKWERNVNFEDGGVTVTQDSSLQDGCFTY 120
```



[illegible]

## RESULT 8

```

US-10-315-920-2
; Sequence 2, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Terbkikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PRO
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discoboma sp.
US-10-315-920-2

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

## RESULT 9

```

; Sequence 4, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic
; TITLE OF INVENTION: Fluorescence
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US//10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT

```

```

; ORGANISM: Discosoma sp.
;
; FEATURE:
;
; OTHER INFORMATION: red fluorescent protein (dsRED)
;
US-10-132-067-4

```

```
Query Match      100.0%; Score 1214; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

## RESULT 10

```

US-10-370-570-56
; Sequence 56, Application US/10370570
; Publication No. US20030219717A1
; GENERAL INFORMATION:
; APPLICANT: DAHL, Soren Weis et al.
; TITLE OF INVENTION: FLUOROPHORE COMPLEMENTATION PRODUCTS
; FILE REFERENCE: 3759-0130P
; CURRENT APPLICATION NUMBER: US/10/370,570
; CURRENT FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Dicosoma sp.
US-10-370-570-56

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	MRSSKNV	I	K	E	F	M	R	F	K	V	R	E	G	T	N	G	H	E	F	E	I	E	G	E	G	R	P	Y	E	G	H	N	T	V	K	L	K	V	T	K	G	E	L	P	F	A	M	I	60				
Db	1	MRSSKNV	I	K	E	F	M	R	F	K	V	R	E	G	T	N	G	H	E	F	E	I	E	G	E	G	R	P	Y	E	G	H	N	T	V	K	L	K	V	T	K	G	E	L	P	F	A	M	I	60				
QY	61	LSPOFOY	G	S	K	V	Y	V	K	H	P	A	D	I	P	D	Y	K	L	S	F	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	V	T	V	T	O	D	S	L	O	D	G	C	F	I	120		
Db	61	LSPOFOY	G	S	K	V	Y	V	K	H	P	A	D	I	P	D	Y	K	L	S	F	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	V	T	V	T	O	D	S	L	O	D	G	C	F	I	120		
QY	121	KVKEFIG	V	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	E	A	S	T	E	R	L	Y	P	R	D	G	V	L	K	G	E	I	H	K	A	L	K	L	K	D	G	H	Y	L	V	E	F	K	S	I	180
Db	121	KVKEFIG	V	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	E	A	S	T	E	R	L	Y	P	R	D	G	V	L	K	G	E	I	H	K	A	L	K	L	K	D	G	H	Y	L	V	E	F	K	S	I	180
QY	181	YMAKKP	V	O	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225														
Db	181	YMAKKP	V	O	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225														
QY	181	YMAKKP	V	O	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225														
Db	181	YMAKKP	V	O	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225														

## RESULT 11

US-10-406-618-32  
; Sequence 32, Application US/10406618  
; Publication No. US20030219814A1  
; GENERAL INFORMATION:  
; APPLICANT: Wan, David Chi-Cheong  
; APPLICANT: Ip, Dennis Tsz-Ming  
; APPLICANT: The Chinese University of Hong Kong



```

; TITLE OF INVENTION: No. US20030219814A1el Fluorescent Proteins
; FILE REFERENCE: 016285-34-1
; CURRENT APPLICATION NUMBER: US/10/406,618
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: US 60/387,968
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US 60/370,598
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Spontaneously
; US-10-406-618-32

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

[illegible]

Qy 61 LSPQFOYGSKYVKHPADIPDYKKLSPEEGFKMERVMNFEDGGVTVTQDSSLQDGCFIY 120

Db 61 LSPQFOYGSKYVKHPADIPDYKKLSPEEGFKMERVMNFEDGGVTVTQDSSLQDGCFIY 120

Qy 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERTLPRDGLKGEIHKALKKDGGHYLVFEKSI 180

Db 121 KYKFIGVNFPSDGPVMQKKTMGWEASTERTLPRDGLKGEIHKALKKDGGHYLVFEKSI 180

Qy	181	YMAKKPVQLPGYVVDSKLDITSHNEDYTIVEQYERTEGRHLLFL	225
Db	181	YMAKKPVQLPGYVVDSKLDITSHNEDYTIVEQYERTEGRHLLFL	225

```

RESULT 12
US-10-433-640-13
; Sequence 13, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Hella
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-433-640-13

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225, Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

cy 1 MRSSKQV I K E F M R F K V R M E G T I V N G H E F E I E G E G E G R P E I G H N T V K L K V T K G G P L P F A W D I 60

Db 1 MRSSKQV I K E F M R F K V R M E G T I V N G H B F E I E G E G E G R P E I G H N T V K L K V T K G G P L P F A W D I 600

61 LSPQFQYSKTVVKHPADIPDYKCLSPQEFKMERVMNFEDGCVTVTQDSSLQDGEFIY 120

Db 61 LSPQFGYSGSKYIVYVHKHPADIPDYKCLSPFEGFKMERVMNFEDGGVTVITQDSSLQDGCFTY 120

QY 121 KVKFIGVNFPSDGPVMQKKTGMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFVKSI 180

Db 121 KVKFIGVNFPSDGPVMQKKTGMGEASTERLYPRDGVLKGEIHKALKLKDGGHYLVFVKSI 180

QY 181 YMAKPPVQLPGYYIVVDSKLDITSHNEDYTVIEQYERTEGRHHLFL 225

Db 181 YMAKPPVQLPGYYIVVDSKLDITSHNEDYTVIEQYERTEGRHHLFL 225

RESULT 13  
US-10-311-030-7

```

: Sequence 7, Application US/10311030
: Publication No. US20040171107A1
: GENERAL INFORMATION:
:   APPLICANT: Nelson, David
:   APPLICANT: Zamiera, Elize
:   APPLICANT: Tselen, Roger
:   TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
:   FILE REFERENCE: 15916-032US1
:   CURRENT APPLICATION NUMBER: US/10/311,030
:   CURRENT FILING DATE: 2002-12-10
:   PRIOR APPLICATION NUMBER: PCT/US01/04625
:   PRIOR FILING DATE: 2001-02-13
:   PRIOR APPLICATION NUMBER: US 60/184,732
:   PRIOR FILING DATE: 2000-02-23
:   NUMBER OF SEQ ID NOS: 12
:   SOFTWARE: FastSeq for Windows Version 4.0
:   SEQ ID NO 7
:   LENGTH: 225
:   TYPE: PRT
:   ORGANISM: Artificial Sequence
:   FEATURE:
:   OTHER INFORMATION: synthetic construct
US-10-311-030-7

```

Query Match	100.0%;	Score 1214;	DB 4;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRSSKAVI	KEFM	RFKV	RM	EG	TV	NG	HE	PI	EG	EG	EG	RP	YE	GH	NT	V	K	L	K	V	T	KG	GP	L	PF	AM	D	I	60
																															60
Db	1	MRSSKAVI	KEFM	RFKV	RM	EG	TV	NG	HE	PI	EG	EG	EG	RP	YE	GH	NT	V	K	L	K	V	T	KG	GP	L	PF	AM	D	I	60

	61	LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFIY	120
Qy			
	61	LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFIY	120
Db			
	61	LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVNNFEDGGVTVTQDSSLQDGCFIY	120

[illegible]

```
OY      181 YMAKPPVQLPGYYYVD SKLDITSHNEDYTIIVEQEYRTGRRHLLFL 225
        |||||
Db       181 YMAKKPVQLPGYYYVD SKLDITSHNEDYTIIVEQEYRTEGRHLLFL 225
```

RESULT 14  
US-10-845-484-3  
; Sequence 3, Application US/10845484  
; Publication No. US20040248180A1  
; AUTOCALCULATION

```

; APPLICANT: Bulina, Maria
; APPLICANT: Chudakov, Dmitry
; APPLICANT: Lukyanov, Konstantin
; TITLE OF INVENTION: Mutant Chromaphores/Fluorophores and
; TITLE OF INVENTION: Methods for Making and Using the Same
; FILE REFERENCE: CLON 092
; CURRENT APPLICATION NUMBER: US/10/845,484
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/343128
; PRIOR FILING DATE: 2001-12-26
;

```

;  
; PRIOR APPLICATION NUMBER: 60/343128  
;  
; PRIOR FILING DATE: 2001-12-26







	Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
	Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
	Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAMDI       				60
Db	1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAMDI       				60
QY	61 LSPQFOYGSKVYVKHPADI PDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFIY       				120
Db	61 LSPQFOYGSKVYVKHPADI PDYKLSFPEGFKMERVMNFDGGVTVTQDSSLQDGCFIY       				120
QY	121 KVKFIGVNFPSDGPVMQCKTMGWEASTERLYPRDGYLKGELHKALKLDGGHYLVEFKSI       				180
Db	121 KVKFIGVNFPSDGPVMQCKTMGWEASTERLYPRDGYLKGELHKALKLDGGHYLVEFKSI       				180
QY	181 YMAKKPVQLPGYYIVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL       				225
Db	181 YMAKKPVQLPGYYIVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL       				225

```

RESULT 18
US-10-505-486-26
/ Sequence 26, Application US/10505486
/ Publication No. US20050118639A1
/ GENERAL INFORMATION:
/ APPLICANT: Takeda Chemical Industries, Ltd.
/ TITLE OF INVENTION: Determination of a ligand
/ FILE REFERENCE: P03-0006PCT
/ CURRENT APPLICATION NUMBER: US/10/505,486
/ CURRENT FILING DATE: 2004-08-20
/ PRIOR APPLICATION NUMBER: JP 2002-45728
/ PRIOR FILING DATE: 2002-02-22
/ PRIOR APPLICATION NUMBER: JP 2002-213949
/ PRIOR FILING DATE: 2002-07-23
/ PRIOR APPLICATION NUMBER: JP 2002-298237
/ PRIOR FILING DATE: 2002-10-11
/ NUMBER OF SEQ ID NOS: 233
/ SEQ ID NO 26
/ LENGTH: 225
/ TYPE: PRT
/ ORGANISM: Dictyosoma sp.
US-10-505-486-26

```

	Query Match	100.0%;	Score 1214;	DB 5;	Length 225;	
	Best Local Similarity	100.0%;	Pred. No. 6.6e-114;			
	Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	
QY	1	MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVTKLKYTKGAPLPFAWDI	60			
DB	1	MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVTKLKYTKGAPLPFAWDI	60			
QY	61	LSPOFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGAVTVTQDSSLQDGCFTY	120			
DB	61	LSPOFOYGSKVYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGAVTVTQDSSLQDGCFTY	120			
QY	121	KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI	180			
DB	121	KYKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI	180			
QY	181	YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225			
DB	181	YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225			

RESULT 19  
US-10-844-064A-2  
; Sequence 2, Application US/10844064A  
; Publication No. US2005014994A1  
; GENERAL INFORMATION:  
; APPLICANT: Bevis, Brooke  
; APPLICANT: Glick, Benjamin  
; TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S

```

; FILE REFERENCE: 092234-9006
; CURRENT APPLICATION NUMBER: US/10/844,064A
; CURRENT FILING DATE: 2004-05-11
; PRIOR APPLICATION NUMBER: PCT/US02/40539
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US 60/341,723
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma
US-10-844-064A-2

Query Match      100.0%; Score 1214; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 6; 6e-114;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

	Query Match	100.0%;	Score 1214;	DB 5;	Length 225;	
	Best Local Similarity	100.0%;	Pred. No. 6.6e-114;			
	Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI					60
Db	1 MRSSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI					60
QY	61 LSPQFOYGSKVYVGHADIPDYKCLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFIY					120
Db	61 LSPQFOYGSKVYVGHADIPDYKQLSPPEGFKMERVMNFEDGGVTVTQDSSLQDGCFIY					120
QY	121 KVKFIGVNFPDSGPVMQKKTMGWBASTERLYPRDGYLKGELIKALKDKDGHYLVEPKSI					180
Db	121 KVKFIGVNFPDSGPVMQKKTMGWEASTERLYPRDGYLKGELIKALKDKDGHYLVEFKSI					180
QY	181 YMAKKPVOLPGYYYVDSKLDITSHNEDYTIVEOYERTEGRHHLPFL					225
Db	181 YMAKKPVOLPGYYYVDSKLDITSHNEDYTIVEOYERTEGRHHLPFL					225

```

RESULT 20
US-10-931-304-1
; Sequence 1, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2Cp3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(225)
; OTHER INFORMATION: wild-type Dsred
US-10-931-304-1

```

Query Match	100.0%;	Score 1214;	DB 5;	Length 225;
Best Local Similarity	100.0%;	Pred. No. 6.6e-114;		
Matches 225; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0



Qy	1	MRSSKNV	IK	EFMR	FK	VR	MEGT	IV	NG	HE	PE	I	B	E	G	E	G	R	P	Y	E	G	H	N	T	V	K	L	K	V	T	K	G	G	P	L	P	F	A	M	I	60												
Db	1	MRSSKNV	IK	EFMR	FK	VR	MEGT	IV	NG	HE	PE	I	B	E	G	E	G	R	P	Y	E	G	H	N	T	V	K	L	K	V	T	K	G	G	P	L	P	F	A	M	I	60												
Qy	61	LSPQFQY	G	S	K	V	Y	V	K	H	P	A	D	I	P	D	Y	K	K	L	S	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	I	O	D	G	C	F	I	120			
Db	61	LSPQFQY	G	S	K	V	Y	V	K	H	P	A	D	I	P	D	Y	K	K	L	S	P	E	G	F	K	M	E	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	I	O	D	G	C	F	I	120			
Qy	121	KVKEIGV	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	W	E	A	S	T	E	R	L	Y	P	R	D	G	V	L	K	G	E	I	H	K	A	L	K	L	K	D	G	H	V	L	V	E	F	K	S	I	180
Db	121	KVKEIGV	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	W	E	A	S	T	E	R	L	Y	P	R	D	G	V	L	K	G	E	I	H	K	A	L	K	L	K	D	G	H	V	L	V	E	F	K	S	I	180
Qy	181	YMAKKP	V	Q	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225															
Db	181	YMAKKP	V	Q	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225															
Qy	181	YMAKKP	V	Q	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225															
Db	181	YMAKKP	V	Q	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	T	I	V	E	Q	Y	E	R	T	E	G	R	H	L	F	L	225															

```

RESULT 21
US-10-343-977-1
; Sequence 1, Application US/10343977
; Publication No. US20040014134A1
; GENERAL INFORMATION:
; APPLICANT: Kuhlemann, Rene
; APPLICANT: Koltermann, Andre
; APPLICANT: Ketting, Ulrich
; APPLICANT: Schwille, Petra
; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/ml
; CURRENT APPLICATION NUMBER: US/10/343,977
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: PCT/EP01/09112
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discosoma sp.
US-10-343-977-1

```

	Query Match	100.0%;	Score 1214;	DB 4;	Length 487;	
	Beet Local Similarity	100.0%;	Pred. No. 1.9e-113;			
	Matches 225;	Conservative	0;	Mismatches	0;	Indels
					Gaps	0;
Qy	1 MRSSKNVIKEFMRFKVMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAMDI       					60
Dd	263 MRSSKNVIKEFMRFKVNMGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTKGGLPFAMDI       					322
Qy	61 LSPQFGYGSKVYVKHPADIPDYKKLSFPGEFKMERVMNFEDGVTVTQTDSLQDGCIFY       					120
Dd	323 LSPQFGYGSKVYVKHPADIPDYKKLSFPGEFKMERVMNFEDGVTVTQTDSLQDGCIFY       					382
Qy	121 KVKFIGVNFPSDGPVMQKKTWGWEASTERLYPRDGVLKGETHKALKDKGGHYLVPEFSI       					180
Dd	383 KVKFIGVNFPSDGPVMQKKTWGWEASTERLYPRDGVLKGETHKALKDKGGHYLVPEFSI       					442
Qy	181 YMAKPPVOLPGYYYVDSKLDITSHNEDIYTIVEQYERTEGRHHLLFL       					225
Dd	443 YMAKPPVOLPGYYYVDSKLDITSHNEDIYTIVEQYERTEGRHHLLFL       					487

RESULT 22  
US-10-343-977-1  
; Sequence 1, Application US/10343977  
; Publication No. US20050112682A9  
; GENERAL INFORMATION:  
; APPLICANT: Kuhlemann, Rene  
; APPLICANT: Kottelmann, Andre  
; APPLICANT: Kottelring, Ulrich  
; APPLICANT: Schwille, Petra

```

; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/ml
; CURRENT APPLICATION NUMBER: US/10/343,977
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: PCT/EP01/09112
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.
US-10-343-977-1

```

[illegible]

```

RESULT 23
US-10-343-977-2
; Sequence 2, Application US/103433977
; Publication No. US20040014134A1
; GENERAL INFORMATION:
; APPLICANT: Kuhlemann, Rene
; APPLICANT: Koltermann, Andre
; APPLICANT: Ketting, Ulrich
; APPLICANT: Schwille, Petra
; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/m1
; CURRENT APPLICATION NUMBER: US/10/343,977
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: PCT/EP01/09112
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.
US-10-343-977-2

```

	Query Match	100.0%;	Score 1214;	DB 4;	Length 506;			
	Best Local Similarity	100.0%;	Pred. No. 2e-113;					
	Matches 225;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0			
Qy	1	MRSSKNV	IKFEMRFKVR	MEGTVNGH	FEIEGEGEGRPYEGHNTV	KLKVTYKGGLPFAMDI	60	
Db	282	MRSSKNV	IKFEMRFKVR	MEGTVNGH	FEIEGEGEGRPYEGHNTV	KLKVTYKGGLPFAMDI	341	
Qy	61	LSPQFOY	GSKVYKHP	ADIPDYK	KLSPPEGFKMERV	MNFEDGCV	VTYTDSSLDGCFIY	120



```
Db      342 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 401
      |||||||
Qy      121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
      |||||||
Db      402 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 461
      |||||||
Qy      181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
      |||||||
Db      462 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 506
      |||||||
```

```
RESULT 24
US-10-343-977-2
; Sequence 2, Application US/10343977
; Publication No. US20050112682A9
; GENERAL INFORMATION:
; APPLICANT: Kuhlemann, Rene
; APPLICANT: Koltermann, Andre
; APPLICANT: Kettling, Ulrich
; APPLICANT: Schwille, Petra
; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/ml
; CURRENT APPLICATION NUMBER: US/10/343,977
; PRIOR APPLICATION NUMBER: 2003-02-06
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.
US-10-343-977-2
```

```
Query Match      100.0%; Score 1214; DB 5; Length 506;
Best Local Similarity 100.0%; Pred. No. 2e-113;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      |||||||
Db      282 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 341
      |||||||
Qy      61 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
      |||||||
Db      342 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 401
      |||||||
Qy      121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
      |||||||
Db      402 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 461
      |||||||
Qy      181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
      |||||||
Db      462 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 506
      |||||||
```

```
RESULT 25
US-10-343-977-3
; Sequence 3, Application US/10343977
; Publication No. US20040014134A1
; GENERAL INFORMATION:
; APPLICANT: Kuhlemann, Rene
; APPLICANT: Koltermann, Andre
; APPLICANT: Kettling, Ulrich
; APPLICANT: Schwille, Petra
; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/ml
; CURRENT APPLICATION NUMBER: US/10/343,977
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: PCT/EP01/09112
```

```
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.
US-10-343-977-3
```

```
Query Match      100.0%; Score 1214; DB 4; Length 547;
Best Local Similarity 100.0%; Pred. No. 2.2e-113;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      |||||||
Db      323 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 382
      |||||||
Qy      61 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
      |||||||
Db      383 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 442
      |||||||
Qy      121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
      |||||||
Db      443 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 502
      |||||||
Qy      181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
      |||||||
Db      503 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 547
      |||||||
```

```
RESULT 26
US-10-343-977-3
; Sequence 3, Application US/10343977
; Publication No. US20050112682A9
; GENERAL INFORMATION:
; APPLICANT: Kuhlemann, Rene
; APPLICANT: Koltermann, Andre
; APPLICANT: Kettling, Ulrich
; APPLICANT: Schwille, Petra
; TITLE OF INVENTION: Dual coloured fluorimetric protease assay
; FILE REFERENCE: 023425us/JH/ml
; CURRENT APPLICATION NUMBER: US/10/343,977
; PRIOR APPLICATION NUMBER: 2003-02-06
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Autofluorescent fusion protein derived from GFP variants of
; OTHER INFORMATION: e.g. Aequorea victoria and/or Discoma sp.
US-10-343-977-3
```

```
Query Match      100.0%; Score 1214; DB 5; Length 547;
Best Local Similarity 100.0%; Pred. No. 2.2e-113;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      |||||||
Db      323 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 382
      |||||||
Qy      61 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
      |||||||
Db      383 LSPQFGYSKYVVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 442
      |||||||
Qy      121 KVKFIGVNPSPDGPVMQKTMGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180
      |||||||
```



Db 443 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 502  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 503 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 547

RESULT 27  
US-10-311-030-9  
Sequence 9, Application US/10311030  
Publication No. US20040171107A1  
GENERAL INFORMATION:  
APPLICANT: Nelson, David  
APPLICANT: Zamlara, Elize  
APPLICANT: Tsien, Roger  
TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
FILE REFERENCE: 15916-032US1  
CURRENT APPLICATION NUMBER: US/10/311,030  
PRIOR FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: PCT/US01/04625  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: US 60/184,732  
PRIOR FILING DATE: 2000-02-23  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 226  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic construct  
US-10-311-030-9

Query Match 99.7%; Score 1210; DB 4; Length 226;  
Best Local Similarity 99.6%; Pred. No. 1.7e-113;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 60  
:|||||  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 61  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 120  
:|||||  
Db 62 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 121  
QY 121 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
:|||||  
Db 122 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 181  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
:|||||  
Db 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 28  
US-10-152-296-2  
Sequence 2, Application US/10152296  
Publication No. US20030077730A1  
GENERAL INFORMATION:  
APPLICANT: Peelle, Beau  
APPLICANT: Rigel Pharmaceuticals, Incorporated  
TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
FILE REFERENCE: 021044-000110US  
CURRENT APPLICATION NUMBER: US/10/152,296  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: US 60/291,871  
PRIOR FILING DATE: 2001-05-18  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 240  
TYPE: PRT  
ORGANISM: Artificial Sequence

: FEATURE:  
: OTHER INFORMATION: Description of Artificial Sequence:mammalian  
: OTHER INFORMATION: codon-optimized variant (DBRED) of Discosoma sp.  
: OTHER INFORMATION: "red" red fluorescent protein (RFP)  
US-10-152-296-2

Query Match 99.7%; Score 1210; DB 4; Length 240;  
Best Local Similarity 99.6%; Pred. No. 1.8e-113;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 60  
:|||||  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 61  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 120  
:|||||  
Db 62 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 121  
QY 121 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
:|||||  
Db 122 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 181  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
:|||||  
Db 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 29  
US-10-739-656-2  
Sequence 2, Application US/10739656  
Publication No. US20040126850A1  
GENERAL INFORMATION:  
APPLICANT: Peelle, Beau  
APPLICANT: Rigel Pharmaceuticals, Incorporated  
TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
FILE REFERENCE: 021044-000110US  
CURRENT APPLICATION NUMBER: US/10/739,656  
CURRENT FILING DATE: 2003-12-17  
PRIOR APPLICATION NUMBER: US/10/152,296  
PRIOR FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: US 60/291,871  
PRIOR FILING DATE: 2001-05-18  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 240  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:mammalian  
OTHER INFORMATION: codon-optimized variant (DBRED) of Discosoma sp.  
OTHER INFORMATION: "red" red fluorescent protein (RFP)  
US-10-739-656-2

Query Match 99.7%; Score 1210; DB 4; Length 240;  
Best Local Similarity 99.6%; Pred. No. 1.8e-113;  
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 60  
:|||||  
Db 2 VRSSKNVKEFMRFKVRMEGTVNGHFEIEGEGGRPYEGHNTVTKLVTKGGPLPFAWDI 61  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 120  
:|||||  
Db 62 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVTVTQDSSLQDGCFTY 121  
QY 121 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 180  
:|||||  
Db 122 KVKFIGVNFPSDGPVMQKKTGWGWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVEFKSI 181  
QY 181 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
:|||||  
Db 182 YMAKKPVQLPGYYVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 226



```
RESULT 30
US-10-214-932-52
; Sequence 52, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214, 932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-214-932-52

Query Match      99.7%; Score 1210; DB 4; Length 545;
Best Local Similarity 99.6%; Pred. No. 5.6e-113;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      : |||
DB      2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 61

QY      61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 120
      |||
DB      62 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 121

QY      121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
      |||
DB      122 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 181

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225
DB      182 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 226

RESULT 31
US-10-214-932-76
; Sequence 76, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214, 932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-214-932-76

Query Match      99.7%; Score 1210; DB 4; Length 548;
Best Local Similarity 99.6%; Pred. No. 5.7e-113;
Matches 224; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      : |||
DB      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
```

```
DB      2 VRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 61

QY      61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 120
      |||
DB      62 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 121

QY      121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
      |||
DB      122 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 181

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225
DB      182 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 226

RESULT 32
US-10-121-258-20
; Sequence 20, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/10/121, 258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed with 1125R
US-10-121-258-20

Query Match      99.4%; Score 1207; DB 4; Length 225;
Best Local Similarity 99.6%; Pred. No. 3.3e-113;
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60
      |||
DB      1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDI 60

QY      61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 120
      |||
DB      61 LSPQFOYGSKYVYKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTY 120

QY      121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
      |||
DB      121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180

QY      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225
DB      181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIQEYERTEGRHHLFL 225

RESULT 33
US-10-315-920-4
; Sequence 4, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315, 920
```











JOURNAL: Nat. Biotech.  
; VOLUME: 17  
; ISSUE: 10  
; PAGES: 969-73  
; DATE: 1999-  
; RELEVANT RESIDUES: (1)..(226)  
US-10-724-178-12

Query Match 98.5%; Score 1195.5; DB 4; Length 226;  
Best Local Similarity 99.1%; Pred. No. 4.9e-112;  
Matches 224; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
181 YMAKKAPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 226  
Db 181 YMAKKAPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 226

RESULT 40  
US-10-315-920-6  
; Sequence 6, Application US/10315920  
; Publication No. US20030175809A1  
; GENERAL INFORMATION:  
; APPLICANT: Fradkov, Arcady Fedorovich  
; APPLICANT: Terstikh, Alexey  
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
; TITLE OF INVENTION: FOR THEIR USE  
; FILE REFERENCE: CLON-077CIP  
; CURRENT APPLICATION NUMBER: US/10/315,920  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/211,607  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: PCT/US01/19097  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: variant of sequence from Discosoma sp.  
US-10-315-920-6

Query Match 98.1%; Score 1191; DB 4; Length 225;  
Best Local Similarity 97.8%; Pred. No. 1.4e-111;  
Matches 220; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 41  
US-10-442-148A-7  
; Sequence 7, Application US/10442148A  
; Publication No. US20040014242A1  
; GENERAL INFORMATION:  
; APPLICANT: IWAKURA, MASAHIRO  
; APPLICANT: HIROTA, KIYONORI  
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND  
; TITLE OF INVENTION: PROCESS FOR ARRAYING AND IMMOBILIZING PROTEIN USING THE SAME  
; FILE REFERENCE: 04583.0103-00000  
; CURRENT APPLICATION NUMBER: US/10/442,148A  
; CURRENT FILING DATE: 2003-05-21  
; PRIOR APPLICATION NUMBER: JP 2002-148950  
; PRIOR FILING DATE: 2002-05-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence  
US-10-442-148A-7

Query Match 97.7%; Score 1186; DB 4; Length 225;  
Best Local Similarity 97.3%; Pred. No. 4.4e-111;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFGYSKYVVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 42  
US-10-442-148A-8  
; Sequence 8, Application US/10442148A  
; Publication No. US20040014242A1  
; GENERAL INFORMATION:  
; APPLICANT: IWAKURA, MASAHIRO  
; APPLICANT: HIROTA, KIYONORI  
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND  
; TITLE OF INVENTION: PROCESS FOR ARRAYING AND IMMOBILIZING PROTEIN USING THE SAME  
; FILE REFERENCE: 04583.0103-00000  
; CURRENT APPLICATION NUMBER: US/10/442,148A  
; CURRENT FILING DATE: 2003-05-21  
; PRIOR APPLICATION NUMBER: JP 2002-148950  
; PRIOR FILING DATE: 2002-05-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 8  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence  
US-10-442-148A-8

Query Match 97.7%; Score 1186; DB 4; Length 239;



Best Local Similarity 97.3%; Pred. No. 4.8e-111;  
Matches 219; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAWDI 60  
DB 1 MASSENVITEFMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
DB 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 43  
US-10-121-258-4  
; Sequence 4, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "T1"  
US-10-121-258-4

Query Match 95.6%; Score 1160; DB 4; Length 225;  
Best Local Similarity 96.0%; Pred. No. 1.8e-108;  
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAWDI 60  
DB 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
DB 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 44  
US-10-931-304-4  
; Sequence 4, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger

; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "T1"  
US-10-931-304-4

Query Match 95.6%; Score 1160; DB 5; Length 225;  
Best Local Similarity 96.0%; Pred. No. 1.8e-108;  
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNTVYKLKVTKGGLPFAWDI 60  
DB 1 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAWDI 60  
QY 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGCFTY 120  
DB 61 LSPQFGYGSKYVVKHPADIPDYKKLSFPEGFKMERVMNFEDGGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
DB 121 KVKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVFEKSI 180  
QY 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
DB 181 YMAKKPVQLPGYYVVDKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 45  
US-10-121-258-24  
; Sequence 24, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/10/121,258  
; CURRENT FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "T1" with 1125R mutation  
US-10-121-258-24

Query Match 95.0%; Score 1153; DB 4; Length 225;



[illegible]

```

RESULT 46
US-10-931-304-24
; Sequence 24, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Desired polypeptide variant "T1" with I125R mutation
US-10-931-304-24

```

	Query Match	95.0%;	Score 1153;	DB 5;	Length 225;
	Best Local Similarity	95.6%;	Pred. No. 9.4e-108;		
	Matches 215;	Conservative	3;	Mismatches 7;	Indels 0;
				Gaps	0.
QY	1 MRSSKNVIKEFMRFKVRMEGTGNGHEFEIEGEGERPYEGHNTVKLKVTKGGLPFAWDI	60			
Db	1 MASSEDVIKBFMRFKVRMEGSVNGHEFEIEGEGERPYEGTQTAKLKVTKGGPLPFAWDI	60			
QY	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTTQDSSLQDGCFIX	120			
Db	61 LSPQFOYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGVTVTTQDSSLQDGSFIY	120			
QY	121 KVKFIGVNFPSDGPVMQKKTMGWEASTERLYPRDVLKGEIHKLKDKDGHYLVEFKSI	180			
Db	121 KVKFRGVNFPSPDGPVMQKKTMGWEASTERLYPRDVLKGEIHKLKDKDGHYLVVEFKSI	180			
QY	181 YMAKKPVQLPGYYIVDSKLDITSHNEDYTIVEQYERTEGRHHLFL	225			
Db	181 YMAKKPVQLPGYYIVDSKLDITSHNEDYTIVEQYERAEGRHHLLFL	225			

## RESULT 47

```

US-10-423-688A-41
; Sequence 41, Application US/10423688A
; Publication No. US20040078148A1
; GENERAL INFORMATION:
; APPLICANT: Los Alamos National Laboratory
; APPLICANT: Waldo, Geoffrey S.
; TITLE OF INVENTION: Directed Evolution Methods for Improving Polypeptide Folding and
; TITLE OF INVENTION: Solubility and Superfolder Fluorescent Proteins Generated Thereby
; FILE REFERENCE: S-100,608
; CURRENT APPLICATION NUMBER: US/10/423,688A
; CURRENT FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: 10/132,067
; PRIOR FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 41
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-10-423-688A-41

```

	Query Match	93.2%	Score 1132;	DB 4;	Length 225;
	Best Local Similarity	92.4%;	Pred. No. 1.2e-105;		
	Matches 208;	Conservative 7;	Mismatches 10;	Indels 0;	Gaps 0
QY	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLP.FAWDI 60    ::     ::				
Dd	1 MESSEDDVIKEFMRFKVHMESGTVNGHEFEIEGEGGRPYEGTQNVKLKVTKGGLP.FAWDI 60    ::     ::				
QY	61 LSPQFOYGSKVVYKHPADIPDYKKLSFPEGFKMERVMNFEDGGVATVTTQDSSLQDGCFIX 120 				
Dd	61 LSPQFOYGSKVVYKHPADIPDYKKLSFPEGFKMERVMNFEDGGVATVTTQDSSLQDGCLIX 120 				
QY	121 KYKFIGVNFPSDGPVMQCKTMGWAEASTERLYPRDGVLKGEIHKALKLKDGGHYLVBEFKSI 180 				
Dd	121 KYKFIGVNFPSDGPVMQCKTMGWEPSTERLYPRDGVLKGDIIHKALKLKDGGHYLVVDIKSI 180 				
QY	181 YMAKKPVOLPGYYYVDSKLDITSHNEDYTIIVEOYERTEGRRHLFL 225 				
Dd	181 YMAKKPVOLPGYYYVDSKLDITNNHEDYTIIVEOYERAEGRRHLFL 225 				

```

RESULT 48
US-10-121-258-6
; Sequence 6, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Taten, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "dimer2"
US-10-121-258-6

```

Query Match	92.3%;	Score 1121;	DB 4;	Length 226;
Best Local Similarity	92.8%;	Pred. No. 1.6e-104;		
Matches 207; Conservative	7;	Mismatches 9;	Indels 0;	Gaps 0;

Qy 3 SSKNVIKEFMRKVRMEGTVNGHEFEIEGEGEGRPRYEGHNTVKLKVTJGGPLPFAMDILS 62



```

Db      4 SSEDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKVTKGGPLPFAWDILS 63
QY      63 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTYKV 122
Db      64 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGTLYKV 123
QY      123 KFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSIYM 182
Db      124 KFRGTNFPDPGPVMQKTMGWEASTERLYPRDGLKGEIHQALKKDGHYLVEFKTIYM 183
QY      183 AKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      184 AKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 226
```

RESULT 49

```

US-10-931-304-6
; Sequence 6, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dsred polypeptide variant "dimer2"
; US-10-931-304-6
```

```

Query Match      92.3%; Score 1121; DB 5; Length 226;
Best Local Similarity 92.8%; Pred. No. 1.6e-104;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
```

```

QY      3 SSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDILS 62
Db      4 SSEDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKVTKGGPLPFAWDILS 63
QY      63 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTYKV 122
Db      64 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGTLYKV 123
QY      123 KFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSIYM 182
Db      124 KFRGTNFPDPGPVMQKTMGWEASTERLYPRDGLKGEIHQALKKDGHYLVEFKTIYM 183
QY      183 AKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      184 AKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 226
```

```

RESULT 50
US-11-052-001-6
; Sequence 6, Application US/11052001
; Publication No. US2005023356A1
; GENERAL INFORMATION:
```

```

; APPLICANT: JONES, JOSHUA T.
; APPLICANT: HAHN, ANGELA TERESA
; APPLICANT: MEYER, TOBIAS
; TITLE OF INVENTION: LIVE-CELL BIOSENSOR POLYPEPTIDES AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: STAN-385
; CURRENT APPLICATION NUMBER: US/11/052,001
; CURRENT FILING DATE: 2005-02-03
; PRIOR APPLICATION NUMBER: 60/541,297
; PRIOR FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fluorescent fusion protein
; US-11-052-001-6
```

```

Query Match      92.3%; Score 1121; DB 6; Length 602;
Best Local Similarity 92.8%; Pred. No. 6.1e-104;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
```

```

QY      3 SSKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDILS 62
Db      142 SSEDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKVTKGGPLPFAWDILS 201
QY      63 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGCFTYKV 122
Db      202 POFQYGSKAYVKHPADIPDYKLSFPEGFKWERVMNPFEDGGVTVTQDSSLQDGTLYKV 261
QY      123 KFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSIYM 182
Db      262 KFRGTNFPDPGPVMQKTMGWEASTERLYPRDGLKGEIHQALKKDGHYLVEFKTIYM 321
QY      183 AKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
Db      322 AKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 364
```

RESULT 51

```

US-10-006-922-44
; Sequence 44, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Lukyanov, Sergey A.
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1 Chromophores/Fluorophores and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 225
; TYPE: PRT
```















```

RESULT 60
US-10-724-178-16
; Sequence 16, Application US/10724178
; Publication No. US20040137528A1
; GENERAL INFORMATION:
; APPLICANT: Odyssey Thera, Inc.
; APPLICANT: Michnick, Stephen
; APPLICANT: Macdonald, Marnie
; APPLICANT: Lamerdin, Jane
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS
; FILE REFERENCE: ODDY007
; CURRENT APPLICATION NUMBER: US/10/724,178
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: US 60/461,133
; PRIOR FILING DATE: 2003-04-09
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: "MRFP1", an engineered monomeric form of DsRed FP
US-10-724-178-16

```

Query Match	84.1%;	Score 1021;	DB 4;	Length 225;
Best Local Similarity	86.9%;	Pred. No. 1.9e-94;		
Matches 192; Conservative	5;	Mismatches 24;	Indels 0;	Gaps 0;

Qy	1	MRSKKNVKEFMRFKVRMEGTVNGHEBPIEGEGEGRPYEGHNTVKLKVTKGGLPFWADI	60
Dd	1	MASSEDVIKEFMRFKVRMEGVSNGHEFEIEGEGBGRPYEGTQTAKLKVTIKGGPLFPWADI	60
Qy	61	LSPQFOYGSKVYVKHPADIPDYKKLSPEPGFKMERVMNPFEDGVVTVTQDSSLQDGCFIY	120
Dd	61	LSPQFOYGSKAIVKHAPADIPDYKLSPFGSKMERVMNPFEDGVVTVTQDSSLQDGSEFIY	120
Qy	121	KVKFIGVNFPSPDGPVMQKKTGMWEASTERLYPRDVLKGEIHKALKLKDGHVLEVEFSI	180
Dd	121	KVKLRGTNFPSPDGPVMQKKTGMWEASTERYPEDGALLKGEIKMRLKLDGDGHDAEVKTT	180
Qy	181	YMAKKPVQLPGYYYYVDskLDITSHNEDYTIVEOYERTTEGRH	221
Dd	181	YMAKKPVQLPGAYKTDIKLDITSHNBEDYTIVEOYERAEGRH	221

```

RESULT 61
US-10-931-304-8
; Sequence 8, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 225

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "mrFP1"
US-10-931-304-8

```

Query Match	84.1%;	Score 1021;	DB 5;	Length 225;
Best Local Similarity	86.9%;	Pred. No. 1.9e-94;		
Matches 192; Conservative	5;	Mismatches 24;	Indels 0;	Gaps 0;

Qy	1	MRSSKNV	IK	FM	R	FKV	ME	G	T	V	N	G	H	F	E	I	E	G	E	G	R	P	Y	E	G	N	T	V	K	L	K	V	T	G	G	P	L	P	F	A	M	D	I	60											
Db	1	MASS	ED	V	I	K	E	F	M	R	F	K	V	A	R	M	E	G	S	V	N	G	H	F	E	I	E	G	E	G	R	P	Y	E	G	I	Q	T	A	K	L	K	V	T	G	G	P	L	P	F	A	M	D	I	60
Qy	61	LSP	Q	F	O	G	S	K	V	Y	V	K	H	P	A	D	I	P	D	Y	K	L	S	P	E	G	F	K	W	E	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	L	O	D	G	C	F	I	Y	120	
Db	61	LSP	Q	F	O	G	S	K	A	Y	V	K	H	P	A	D	I	E	D	Y	L	K	L	S	P	E	G	F	K	W	E	R	V	M	N	F	E	D	G	V	T	V	T	Q	D	S	L	O	D	G	E	F	I	Y	120
Qy	121	KV	K	F	I	G	V	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	E	A	S	T	E	R	L	Y	P	R	D	G	V	L	K	G	E	I	H	K	L	K	D	G	H	Y	L	V	E	F	K	S	I	180
Db	121	KV	K	L	R	G	T	N	F	P	S	D	G	P	V	M	Q	K	T	M	G	E	A	S	T	E	R	M	Y	P	E	D	G	A	L	K	G	E	I	K	M	R	L	K	D	G	H	Y	D	A	E	V	K	T	180
Qy	181	YMA	K	K	P	V	Q	L	P	G	Y	Y	V	D	S	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	T	E	G	R	H	221															
Db	181	YMA	K	K	P	V	Q	L	P	G	A	Y	K	T	D	I	K	L	D	I	T	S	H	N	E	D	Y	T	I	V	E	Q	Y	E	R	A	E	G	R	H	221														

RESULT 62  
US-10-931-304-108  
; Sequence 108, Application US/10931304  
; Publication No. US20050196768A1

```

; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dered polypeptide variant "mGrapel"
US-10-931-304-108

```

	Query Match	83.5%	Score 1014;	DB 5;	Length 225;
	Best Local Similarity	86.4%;	Pred. No. 9.8e-94;		
	Matches 191; Conservative	6;	Mismatches 24;	Indels 0;	Gaps 0;
QY	1 MRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLKVTGGPLPFAWDI	60			
Dd	1 MASSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGERPYEGTQTAKLVTKGGPLPFAWDI	60			
QY	61 LSPQFGYSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGCVTVTQDSSLQDGCFIY	120			
Dd	61 LSQPFMYGSKAYVKHPADIPYLKLSPPEGFKMERVMNFEDGVTVTQDSSLQDEFIY	120			
QY	121 KYKFIGVNPSPDGPVMQKKTMGWEASTERLYPRDGLKGETHKALKLKDGHYLVFEFSI	180			
Dd	121 KVLIRGTNFPSDGPVMQKKTMGWEASSERLYPEDGALKEIKMRLKLKDGHYDAEAQTT	180			



QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 181 YMAKRPVQLPGAYKLDYKLDITSHNEDYTIIVEQYERABEGRH 221

RESULT 63  
US-10-931-304-79  
; Sequence 79, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 79  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "mRFP1.1"  
US-10-931-304-79

Query Match 83.4%; Score 1012; DB 5; Length 225;  
Best Local Similarity 86.0%; Pred. No. 1.6e-93;  
Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMDI 60  
Db 1 MASSEDVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGTQAKLKVTGGPLPFAMDI 60  
QY 61 LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGEFIY 120  
Db 61 LSPQFMGSKAYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGEFIY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
Db 121 KYKLKGTNFPSPDGPVMQKTMGWEASERMYPEDGALKGEIKRLKLDGGHYDAEVKTT 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 181 YMAKRPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERABEGRH 221

RESULT 64  
US-10-931-304-104  
; Sequence 104, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29

; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 104  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "mHoneydew"  
US-10-931-304-104

Query Match 82.9%; Score 1007; DB 5; Length 225;  
Best Local Similarity 85.1%; Pred. No. 5e-93;  
Matches 188; Conservative 8; Mismatches 25; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMDI 60  
Db 1 MASSEDVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGTQAKLKVTGGPLPFAMDI 60  
QY 61 LSPQFGYGSKVYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGEFIY 120  
Db 61 LSPQFMGSKAYVKHPADIPDYKLSFPEGFKMERVMNFEDEGVTVTQDSSLQDGEFIY 120  
QY 121 KYKFIGVNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180  
Db 121 KYKLKGTNFPSPDGPVMQKTMGWAATTERMYPEDGALKGEIKRLKLDGGHYDAEVKTT 180  
QY 181 YMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 181 YMAKRPVQLPGAYKIDKLDITSHNEDYTIIVEQYERABEGRH 221

RESULT 65  
US-10-931-304-100  
; Sequence 100, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 100  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "mTangerine"  
US-10-931-304-100

Query Match 82.5%; Score 1002; DB 5; Length 225;  
Best Local Similarity 85.5%; Pred. No. 1.6e-92;  
Matches 189; Conservative 6; Mismatches 26; Indels 0; Gaps 0;  
QY 1 MRSSKNVKEFMRFKVRMEGTGNGHEFEIEGEGGRPYEGHNTVKLKVTGGPLPFAMDI 60



```
Db      1 MASSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAMDI 60
QY      61 LSPQFQYGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTY 120
Db      61 LSPQFCYGSKAYVKHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGEFTY 120
QY      121 KYKPIGVNFPSPDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSI 180
Db      121 KYLGRGTNFPSPDGPVMQKTMGWEASSERMYPEDGALKGEIKMRLLKLDGGHYDAEVKTT 180
QY      181 YMAKPEVQLPGYYVDSKLDITSHNEDYTIIVEOYERTEGRH 221
Db      181 YMAKPEVQLPGAYKTDIKLDITSHNEDYTIIVELYERAEGRH 221
```

```
RESULT 66
US-10-931-304-85
; Sequence 85, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 85
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "Ors4-9"
US-10-931-304-85
```

```
Query Match      81.9%; Score 994; DB 5; Length 236;
Best Local Similarity 86.5%; Pred. No. 1.1e-91;
Matches 186; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY      7 VIKFEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDILSPQFQ 66
Db      12 IIKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAMDILSPQFT 71
QY      67 YGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTYKVKFIG 126
Db      72 YGSKAYVKHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGEFTYKVKLRG 131
QY      127 VNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSIYMAKKP 186
Db      132 TNFPSDGPVMQKTMGWEASSERMYPEDGALKGEIKMRLLKLDGGHYDAEVKTTYKAKKP 191
QY      187 VOLPGYYVDSKLDITSHNEDYTIIVEOYERTEGRH 221
Db      192 VOLPGAYKVDIKLDITSHNEDYTIIVEOYERAEGRH 226
```

```
RESULT 67
US-10-931-304-89
; Sequence 89, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 89
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mFRFP (F2Q6) (mGrape2)"
US-10-931-304-89
```

```
Query Match      81.7%; Score 992; DB 5; Length 236;
Best Local Similarity 86.0%; Pred. No. 1.7e-91;
Matches 185; Conservative 6; Mismatches 24; Indels 0; Gaps 0;

QY      7 VIKFEMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAMDILSPQFQ 66
Db      12 VIKFEMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAMDILSPQFM 71
QY      67 YGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGCFTYKVKFIG 126
Db      72 YGSKAYVKHPDIPDYKLSFPEGFKMERVMNPFEDGVVTVTQDSSLQDGEFTYKVKLRG 131
QY      127 VNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLDGGHYLVEFKSIYMAKKP 186
Db      132 TNFPSDGPVMQKTMGWEASSERLYPEDGALKGEIKMRLLKLDGGHYDAEVKTTYMAKKP 191
QY      187 VOLPGYYVDSKLDITSHNEDYTIIVEOYERTEGRH 221
Db      192 VOLPGAYKLDYKLDITSHNEDYTIIVEOYERAEGRH 226
```

```
RESULT 68
US-10-931-304-102
; Sequence 102, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 102
; LENGTH: 236
; TYPE: PRT
```



ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dsred polypeptide variant "mOrange (mOFp1)"  
US-10-931-304-102

Query Match 81.5%; Score 989; DB 5; Length 236;  
Best Local Similarity 86.0%; Pred. No. 3.5e-91;  
Matches 185; Conservative 5; Mismatches 25; Indels 0; Gaps 0;  
QY 7 VIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKYTKGGPLPFAWDILSPQFQ 66  
:|||||||:|||||||  
Db 12 IIKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGFQTAKLKYTKGGPLPFAWDILSPQFT 71  
QY 67 YGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTYKYKFIG 126  
|||||||:|||||||  
Db 72 YGSKAYVKHPADIPDYFKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGEFTYKYLRG 131  
QY 127 VNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186  
:|||||||:|||||||  
Db 132 TNFPSDGPVMQKTMGWEASSERMPEDGALKGEIKORLKLKDGGHYTSEVKTTYKAKKP 191  
QY 187 VQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 192 VQLPGAYIVGIKLDITSHNEDYTIIVEQYERAEGRH 226

RESULT 69  
US-10-931-304-83  
; Sequence 83, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 83  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dsred polypeptide variant "mRFP1.5"  
US-10-931-304-83

Query Match 81.4%; Score 988; DB 5; Length 236;  
Best Local Similarity 85.6%; Pred. No. 4.4e-91;  
Matches 184; Conservative 6; Mismatches 25; Indels 0; Gaps 0;  
QY 7 VIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKYTKGGPLPFAWDILSPQFQ 66  
:|||||||:|||||||  
Db 12 IIKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKYTKGGPLPFAWDILSPQFM 71  
QY 67 YGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTYKYKFIG 126  
|||||||:|||||||  
Db 72 YGSKAYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGEFTYKYLRG 131  
QY 127 VNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186  
:|||||||:|||||||  
Db 132 TNFPSDGPVMQKTMGWEASSERMPEDGALKGEIKORLKLKDGGHYDAEVKTTYKAKKP 191

QY 187 VQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
|||||||:|||||||  
Db 192 VQLPGAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 226

RESULT 70  
US-10-931-304-92  
; Sequence 92, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 92  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Dsred polypeptide variant "mRFP2 (mCherry)"  
US-10-931-304-92

Query Match 81.3%; Score 987; DB 5; Length 236;  
Best Local Similarity 85.6%; Pred. No. 5.6e-91;  
Matches 184; Conservative 7; Mismatches 24; Indels 0; Gaps 0;  
QY 7 VIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKYTKGGPLPFAWDILSPQFQ 66  
:|||||||:|||||||  
Db 12 IIKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKYTKGGPLPFAWDILSPQFM 71  
QY 67 YGSKVYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGCFTYKYKFIG 126  
|||||||:|||||||  
Db 72 YGSKAYVKHPADIPDYKLSFPEGFKMERVMNPFEDGGVVTYTQDSSLQDGEFTYKYLRG 131  
QY 127 VNFPSDGPVMQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186  
:|||||||:|||||||  
Db 132 TNFPSDGPVMQKTMGWEASSERMPEDGALKGEIKORLKLKDGGHYDAEVKTTYKAKKP 191  
QY 187 VQLPGYVYVDSKLDITSHNEDYTIIVEQYERTEGRH 221  
Db 192 VQLPGAYNVNIKLDITSHNEDYTIIVEQYERAEGRH 226

RESULT 71  
US-10-931-304-94  
; Sequence 94, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258



```

; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 94
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "mYOFp (74-11)"
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 62
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-931-304-94
```

```

Query Match      81.2%; Score 986; DB 5; Length 235;
Best Local Similarity 85.6%; Pred. No. 7e-91;
Matches 184; Conservative 5; Mismatches 26; Indels 0; Gaps 0;
```

```

QY 7 VIKFPMRFKVRMEGTVNGHFEIEGEGEGRPYEGHNTVTKLVTKGGLPFAMWILSPQFQ 66
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIKFPMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTAKLVTKGGLPFAMWILSPQFT 71

QY 67 YGSKVYVKHPADIPDYKLLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGCFTYKVKFIG 126
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAYVKHPADIPDYKLKLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGEFTYKVKLRG 131

QY 127 VNFPSDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKTMGWEASSERMPEDGALKGEIKRLKLDKGHYTSEVKTYYKAKKP 191

QY 187 VOLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRH 221
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTYVEQYERAEGRH 226
```

```

RESULT 72
US-10-931-304-96
; Sequence 96, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 96
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "mROFP (A2/6-6)"
US-10-931-304-96
```

```

Query Match      80.1%; Score 973; DB 5; Length 236;
Best Local Similarity 84.7%; Pred. No. 1.4e-89;
```

```

Matches 182; Conservative 6; Mismatches 27; Indels 0; Gaps 0;

QY 7 VIKFPMRFKVRMEGTVNGHFEIEGEGEGRPYEGHNTVTKLVTKGGLPFAMWILSPQFQ 66
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIKFPMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQAKLVTKGGLPFAMWILTPNFT 71

QY 67 YGSKVYVKHPADIPDYKLLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGCFTYKVKFIG 126
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAYVKHPADIPDYKLKLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGEFTYKVKLRG 131

QY 127 VNFPSDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKTMGWEASSERMPEDGALKGEIKRLKLDKGHYDAEVKTTYKAKKP 191

QY 187 VOLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRH 221
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTYVEQYERAEGRH 226
```

```

RESULT 73
US-10-931-304-98
; Sequence 98, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; PRIOR FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 98
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed polypeptide variant "mStrawberry"
US-10-931-304-98
```

```

Query Match      80.1%; Score 973; DB 5; Length 236;
Best Local Similarity 84.7%; Pred. No. 1.4e-89;
Matches 182; Conservative 6; Mismatches 27; Indels 0; Gaps 0;
```

```

QY 7 VIKFPMRFKVRMEGTVNGHFEIEGEGEGRPYEGHNTVTKLVTKGGLPFAMWILSPQFQ 66
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 IIKFPMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQAKLVTKGGLPFAMWILTPNFT 71

QY 67 YGSKVYVKHPADIPDYKLLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGCFTYKVKFIG 126
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 YGSKAYVKHPADIPDYKLKLSFPEGFKWERVWVNFEDGVTVTQDSSLQDGEFTYKVKLRG 131

QY 127 VNFPSDGPVWQKTMGWEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSIYMAKKP 186
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 132 TNFPSDGPVWQKTMGWEASSERMPEDGALKGEIKRLKLDKGHYDAEVKTTYKAKKP 191

QY 187 VOLPGYVYVDSKLDITSHNEDYTYVEQYERTEGRH 221
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 192 VOLPGAYIVGIKLDITSHNEDYTYVEQYERAEGRH 226
```

```

RESULT 74
US-10-931-304-87
```



```

; Sequence 87, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dated polypeptide variant "Y1.3 (MYOP 1.3) (mBanana)"
US-10-931-304-87

```

	Query Match	79.1%;	Score 960;	DB 5;	length 236;	
	Best Local Similarity	83.7%;	Pred. No. 2.9e-88;			
	Matches 180;	Conservative	5;	Mismatches 30;	Indels 0;	Gaps 0;
Qy	7 VIKFEMRFKVRMEGTVNGHFEFEIEGEGERPYEGHNTVKLKVTKGGLPFAMDILSPQFC		:			
Dd	12 VIKFEMRFKVRMEGSVNCHFEFEIEGEGERPYEGTQTAKLKVTKGGPLPFAMDILSPQFC		:			
Qy	67 YGSKVVYVKHPADI PDYKKLSFPFGFKMERVMNFEDGGVTVYTQDSSLQDGCFITYKKVFIS		:			
Dd	72 YGSKAYYVKHPGTGI PDYFKLSFPEGFKMERVMNFFEDGGVTVYAQDSSLQDGEFIYKKVLRG		:			
Qy	127 VNFPSDGPVMOCKKTGMWEASTERLYPRDGVLKGEIHKALKLKDGGHYLVVEFKSIYMAKP		:			
Dd	132 TNFPSDGPVMOCKKTGMWEASSERMY PEDGALKEIKMRLLKLKDGGHYSAETKTTYKAKRP		:			
Qy	187 VOLPGYYYVD SKLDITSHNEDYTIIVEOYERTEGRH		:			
Dd	192 VOLPGAYIAGEKIDITSHNEDYTIIVELYERAEGRH		:			

```

RESULT 75
US-10-724-178-1041
: Sequence 1041, Application US/10724178
: Publication No. US20040137528A1
: GENERAL INFORMATION:
: APPLICANT: Odyssey Thera, Inc.
: APPLICANT: Michnick, Stephen
: APPLICANT: Macdonald, Marnie
: APPLICANT: Lamerdin, Jane
: TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
: TITLE OF INVENTION: COMPLEMENTATION ASSAYS
: FILE REFERENCE: ODDY007
: CURRENT APPLICATION NUMBER: US/10/724,178
: CURRENT FILING DATE: 2003-12-01
: PRIOR APPLICATION NUMBER: US 60/461,133
: PRIOR FILING DATE: 2003-04-09
: NUMBER OF SEQ ID NOS: 1067
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1041
: LENGTH: 183
: TYPE: PRT
: ORGANISM: Artificial
: FEATURE:

```

```

; OTHER INFORMATION: modif. frag.; RFP file, with position 1 Met removed
US-10-724-178-1041

Query Match      69.9%; Score 848; DB 4; Length 183;
Best Local Similarity 87.4%; Pred. No. 4, 1e-77;
Matches 159; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY      3 SSKNVIKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFPAMDILS 62
      ||:|||||
Db      2 SSBDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQTAKLKVTKGGLPFPAMDILS 61

QY      63 PQFOYGSKYVYKHPADIPDYKKLSFPEGFKMERVMNFFEDGGVTVTQDSSLQDGCFIYKV 122
      |||
Db      62 PQFOYGSKAYVYKHPADIPDYKLSFPEGFKMERVMNFFEDGGVTVTQDSSLQDGEFIYKV 121

QY      123 KFIGVNFPSDGPVMQKKTMGWEASTERLYPRDGLKGEIHKALKKDGGHYLVPEFSIYM 182
      |||
Db      122 KLRGTNFPSDGPVMQKKTMGWEASTERMYPRDGLKGEIKRLLKXDGGHYDAEVKITYM 181

QY      183 AK 184
      ||
Db      182 AK 183

```

Search completed: January 11, 2006, 02:12:06  
Job time : 64 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 11, 2006, 02:08:34 ; Search time 9 Seconds  
(without alignments)  
212.341 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIQEYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 7

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Database : Published Applications AA New:\*

1:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1214	100.0	225	6	US-10-209-208-1
2	1207	99.4	225	6	US-10-209-208-20
3	1160	95.6	225	6	US-10-209-208-4
4	1153	95.0	225	6	US-10-209-208-24
5	1121	92.3	226	6	US-10-209-208-6
6	1021	84.1	225	6	US-10-209-208-8
7	1012	83.4	225	6	US-10-209-208-79

ALIGNMENTS

RESULT 1  
US-10-209-208-1  
; Sequence 1, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29

PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(225)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-1

Query Match 100.0%; Score 1214; DB 6; Length 225;  
Best Local Similarity 100.0%; Pred. No. 2.3e-110;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTGNGHFEIEGEGEPYEGHNTVKLKTKGGLPFAWDI	60
DB	1	MRSSKNVKEFMRFKVRMEGTGNGHFEIEGEGEPYEGHNTVKLKTKGGLPFAWDI	60
QY	61	LSPPQFGSKVYVYKHPADIPYKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCIFY	120
DB	61	LSPPQFGSKVYVYKHPADIPYKLSFPEGFKWERVMNFDGGVVTVTQDSSLQDGCIFY	120
QY	121	KVKFIGNVPPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKSI	180
DB	121	KVKFIGNVPPSDGPVMQKKTGWEASTERLYPRDGLKGEIHKALKDKGHYLVEFKSI	180
QY	181	YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225
DB	181	YMAKKPVQLPGYVVYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL	225

RESULT 2  
US-10-209-208-20  
; Sequence 20, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed with I125R  
US-10-209-208-20

Query Match 99.4%; Score 1207; DB 6; Length 225;  
Best Local Similarity 99.6%; Pred. No. 1.1e-109;  
Matches 224; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MRSSKNVKEFMRFKVRMEGTGNGHFEIEGEGEGRPYEGHNTVKLKTKGGLPFAWDI	60
----	---	--	----



Db 1 MRSKKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFRGVNFPDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 3  
US-10-209-208-4  
; Sequence 4, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey, Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "T1"  
US-10-209-208-4

Query Match 95.6%; Score 1160; DB 6; Length 225;  
Best Local Similarity 96.0%; Pred. No. 3.7e-105;  
Matches 216; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MRSKKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSEVDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 4  
US-10-209-208-24  
; Sequence 24, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert

; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24  
; LENGTH: 225  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "T1" with I125R mutation  
US-10-209-208-24

Query Match 95.0%; Score 1153; DB 6; Length 225;  
Best Local Similarity 95.6%; Pred. No. 1.8e-104;  
Matches 215; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MRSKKNVKEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTVKLKVTKGGLPFAWDI 60  
Db 1 MASSEVDVKEFMRFKVRMEGSVNGHEFEIEGEGGRPYEGTQAKLKVTKGGLPFAWDI 60  
QY 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120  
Db 61 LSPQFQYGSKVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGSFTY 120  
QY 121 KVKFIGVNFPSDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
Db 121 KVKFRGVNFPDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVEFKSI 180  
QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225  
Db 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225

RESULT 5  
US-10-209-208-6  
; Sequence 6, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey, Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: DsRed polypeptide variant "dimer2"  
US-10-209-208-6

Query Match 92.3%; Score 1121; DB 6; Length 226;



```

Best Local Similarity 92.8%; Pred. No. 2.2e-101;
Matches 207; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 3 SSKNVIKEMRPFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKVTKGGLPFAMWIDLS 62
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 SSEVDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQTAKLKVTKGGLPFAMWIDLS 63

QY 63 PQFGYGSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTYKV 122
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 PQFGYGSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGLTYKV 123

QY 123 KPIGVNFPSPDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSIYM 182
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 124 KFRGTNFPDPGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKTIYM 183

QY 183 AKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHHLFL 225
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 AKKPVQLPGYYVDTKLDITSHNEDYTIIVEQYERSEGRHHLFL 226

```

RESULT 6

```

US-10-209-208-8
; Sequence 8, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mrFP1"
US-10-209-208-8

```

```

Query Match 84.1%; Score 1021; DB 6; Length 225;
Best Local Similarity 86.9%; Pred. No. 1e-91;
Matches 192; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY 1 MRSSKNVIEFMRPFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKVTKGGLPFAMWI 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MASSEDVIEFMRPFKVRMEGSVNGHEFEIEGEGRPYEGTQTAKLKVTKGGLPFAMWI 60

QY 61 LSPQFGYGSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LSPQFGYGSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGEFTY 120

QY 121 KVKFIGVNFPSPDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 KVKLRGTNFPSPDGPVMQKKTGMWEASTERMYPEDGALKGEIKRLKLDGHHYDAEVKTT 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 YMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 221

```

```

RESULT 7
US-10-209-208-79
; Sequence 79, Application US/10209208

```

```

; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Dared polypeptide variant "mrFP1.1"
US-10-209-208-79

```

```

Query Match 83.4%; Score 1012; DB 6; Length 225;
Best Local Similarity 86.0%; Pred. No. 7.4e-91;
Matches 190; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 1 MRSSKNVIEFMRPFKVRMEGTVNGHEFEIEGEGRPYEGHNTVTKVTKGGLPFAMWI 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MASSEDVIEFMRPFKVRMEGSVNGHEFEIEGEGRPYEGTQTAKLKVTKGGLPFAMWI 60

QY 61 LSPQFGYGSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGCFTY 120
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LSPQFMYSKAVYVKHPADIPDYKKLSFPEGFKMERVMNPFEDGGVTVTQDSSLQDGEFTY 120

QY 121 KVKFIGVNFPSPDGPVMQKKTGMWEASTERLYPRDGLKGEIHKALKKDGHYLVEFKSI 180
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 KVKLRGTNFPSPDGPVMQKKTGMWEASSERMYPEDGALKGEIKRLKLDGHHYDAEVKTT 180

QY 181 YMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRH 221
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 YMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAEGRH 221

```

```

Search completed: January 11, 2006, 02:12:24
Job time : 10 secs

```



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:23:38 ; Search time 3690 Seconds  
(without alignments)  
3466.064 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 85

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10006922/runat.10012006.162504.19114/app\_query.fasta.1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2nmInscr.rge -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=500 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=69 -ALIGN=500  
-MODE=LOCAL -OUTFMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN 1 1 7415 @runat.10012006.162504.19114 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_scs:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hcg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	6	AR183915 Sequence
2	1214	100.0	678	6	AX172854 Sequence
3	1214	100.0	678	6	AX207715 Sequence

4	1214	100.0	678	6	AX233581	AX233581 Sequence
5	1214	100.0	678	6	AX370404	AX370404 Sequence
6	1214	100.0	678	6	AX824725	AX824725 Sequence
7	1214	100.0	859	2	AF168419	AF168419 Discosoma
8	1214	100.0	859	6	AR636082	AR636082 Sequence
9	1214	100.0	859	6	AR670157	AR670157 Sequence
10	1214	100.0	859	6	AX463698	AX463698 Sequence
11	1210	99.7	681	6	AX233584	AX233584 Sequence
12	1210	99.7	713	6	AX233627	AX233627 Sequence
13	1210	99.7	723	6	AR527331	AR527331 Sequence
14	1210	99.7	2721	6	CQ882115	CQ882115 Sequence
15	1210	99.7	2772	6	CQ882117	CQ882117 Sequence
16	1210	99.7	4692	6	AX463702	AX463702 Sequence
17	1210	99.7	6893	6	AX823860	AX823860 Sequence
18	1210	99.7	8811	11	AY569779	AY569779 Cloning v
19	1210	99.7	9320	6	AX663075	AX663075 Sequence
20	1210	99.7	12404	11	AY569780	AY569780 Cloning v
21	1207	99.4	678	6	AX370406	AX370406 Sequence
22	1196	98.5	666	6	AX348043	AX348043 Sequence
23	1196	98.5	666	6	AX353910	AX353910 Sequence
24	1196	98.5	711	2	AY679107	AY679107 Discosoma
25	1193	98.3	678	2	AF545828	AF545828 Discosoma
26	1191	98.1	678	6	AX370408	AX370408 Sequence
27	1191	98.1	678	6	AX824732	AX824732 Sequence
28	1191	98.1	921	2	AY679106	AY679106 Discosoma
29	1186	97.7	675	6	AX824731	AX824731 Sequence
30	1186	97.7	1050	6	AX666133	AX666133 Sequence
31	1186	97.7	4488	6	CQ849509	CQ849509 Sequence
32	1186	97.7	4546	11	AY818375	AY818375 Cloning v
33	1186	97.7	4555	11	DQ005468	DQ005468 Cloning v
34	1186	97.7	4570	11	AY818373	AY818373 Cloning v
35	1186	97.7	4829	11	AY640628	AY640628 SiRNA vec
36	1186	97.7	4906	11	AY640625	AY640625 SiRNA vec
37	1186	97.7	5311	11	AJ851284	AJ851284 Cloning v
38	1186	97.7	6423	11	AY613997	AY613997 Cloning v
39	1186	97.7	6990	11	AY640633	AY640633 SiRNA vec
40	1186	97.7	7058	11	AY640630	AY640630 SiRNA vec
41	1186	97.7	7147	11	CQ849511	CQ849511 Sequence
42	1186	97.7	7616	6	CS018252	CS018252 Sequence
43	1186	97.7	9468	6	CS018253	CS018253 Sequence
44	1186	97.7	10140	6	CS018254	CS018254 Sequence
45	1186	97.7	10522	6	CS018256	CS018256 Sequence
46	1186	97.7	10786	6	CS018251	CS018251 Sequence
47	1186	97.7	11251	6	CS018255	CS018255 Sequence
48	1186	97.7	11570	6	CS018249	CS018249 Sequence
49	1186	97.7	11867	6	CS018257	CS018257 Sequence
50	1186	97.7	11920	6	CQ981073	CQ981073 Sequence
51	1186	97.7	14720	6	CQ981074	CQ981074 Sequence
52	1186	97.7	16157	6	AX686888	AX686888 Sequence
53	1184	97.5	898	6	AY916793	AY916793 Dual fluo
54	1170	96.4	3441	11	AJ851285	AJ851285 Cloning v
55	1170	96.4	5311	11	AJ851289	AJ851289 Cloning v
56	1170	96.4	13079	11	AJ851286	AJ851286 Cloning v
57	1165	96.0	5311	11	AY342347	AY342347 Red H-Pel
58	1165	96.0	10141	11	AY342348	AY342348 Red H-Sti
59	1165	96.0	10276	11	AY490568	AY490568 UAS-Red S
60	1165	96.0	10481	11	AJ851290	AJ851290 Cloning v
61	1165	96.0	13079	11	AF506025	AF506025 Synthetic
62	1165	96.0	681	11	AF506026	AF506026 Synthetic
63	1121	92.3	1395	11	AX824729	AX824729 Sequence
64	1119	92.2	678	6	AY678268	AY678268 Synthetic
65	1095	90.2	705	11	AF272711	AF272711 Discosoma
66	1095	90.2	1431	11	AX686894	AX686894 Sequence
67	1085.5	89.4	876	2	AF506027	AF506027 Synthetic
68	1085.5	89.4	876	6	AB166761	AB166761 Synthetic
69	1085.5	89.4	876	11	DQ005472	DQ005472 Cloning v
70	1021	84.1	696	11	DQ005474	DQ005474 Cloning v
71	1021	84.1	4555	11	DQ005475	DQ005475 Cloning v
72	1021	84.1	5311	11	AJ851287	AJ851287 Cloning v
73	1021	84.1	13079	11	AJ851291	AJ851291 Cloning v
74	1021	84.1				
75	1021	84.1				
76	1021	84.1				



77	1007	82.9	678	11	AY678271	AY678271 Synthetic
78	1002	82.5	678	11	AY678270	AY678270 Synthetic
79	994	81.9	702	11	AY679163	AY679163 Synthetic
80	989	81.5	711	11	AY678265	AY678265 Synthetic
81	987	81.3	711	11	AY678264	AY678264 Synthetic
82	985	81.1	681	2	AY786536	AY786536 Discosoma
83	973	80.1	711	11	AY678266	AY678266 Synthetic
84	972	80.1	681	2	AY786537	AY786537 Discosoma
85	960	79.1	711	11	AY678267	AY678267 Synthetic

ALIGNMENTS

RESULT 1					
ARI83915					
LOCUS	ARI83915	678 bp	DNA	linear	PAT 20-APR-2002
DEFINITION	Sequence 6 from patent US 6342379.				
ACCESSION	ARI83915				
VERSION	ARI83915.1	GI:20227884			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 678)				
AUTHORS	Tsien,R.Y. and Gonzalez,J.E. III.				
TITLE	Detection of transmembrane potentials by optical methods				
JOURNAL	Patent: US 6342379-A 6 29-JAN-2002;				
FEATURES	location/Qualifiers				
Source	1..678				
	/organism="unknown"				
	/mol_type="unassigned DNA"				

ORIGIN

Alignment Scores:					
Pred. No.:	1.4e-114	Length:	678		
Score:	1214.00	Matches:	225		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		

US-10-006-922A-12 (1-225) x ARI83915 (1-678)

QY	1	MetArgSerSerLySAAnValIleLySGluPheMetArgPheLySValArgMetGluGly	20
DB	1	ATGAGGCTCTCCAGAAGATGTTATCAAGAGTTCATGAGGTTTAAGTTTCGCATGAGACA	60
QY	21	ThrValAAnGLyHISgluPheGluIleGluGlyGluGlyArgProTyrgluGly	40
DB	61	ACGGTCAATGGGACGAGTTTGAATAGAGCGAAGAGGGGAGGCCATACGAAGGC	120
QY	41	HISAnThrValLyLeuLySValThrLySGlyGlyProLeuProPheAlaTPAapIle	60
DB	121	CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrglySerLySValTyrrVallySHISProAlaApIlePro	80
DB	181	TTGTCAACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCCATACCA	240
QY	81	AspTyrrLySlySLeuSerPheProGluGlyPheLySTrpGluArgValMetAAnPheGlu	100
DB	241	GACTATAAAAAAGCTGTCATTTCCTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTyrr	120
DB	301	GACGGTGGCGTCGTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTCACTAAC	360
QY	121	LySValLySpheIleGlyValAAnPheProSerAspGlyProValMetGlnLySlyThr	140
DB	361	AAGGTCAAGTTCATTGGCGGTGAACCTTTCCTCCGATGGACCTGTATGCAAAAAGACACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLySGlyGlu	160

DB	421	ATGGCTGGGAAGCCAGCACTGAGCGTTTGATATCCTCGTATGCGCGTGTGAAGAGAG	480
QY	161	IleHISlySAleuLyLeuLySAAspGlyGlyHISTyrrLeuValGluPheLySerIle	180
DB	481	ATTCTAAGGCTCTGAAGCTGAAGAGCGGTGTCATTACTAGTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLySlySProValGlnLeuProGlyTyrrTyrrTyrrValAAspSerLySLeuA	200
DB	541	TACATGGCAAGAGCGCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAAACTG	600
QY	201	IleThrSerHISAnGluAspTyrrThrIleValGluGlnTyrrGluArgThrGluGlyArg	220
DB	601	ATAACAAGCCACACAGACTATACATCGTTGACGAGTATGAAGAACCAGAGGAGCG	660
QY	221	HISHisLeuPheLeu	225
DB	661	CACCATCTGTTCCTT	675

RESULT 2					
AX172854					
LOCUS	AX172854	678 bp	DNA	linear	PAT 03-JUL-2001
DEFINITION	Sequence 6 from Patent WO0142211.				
ACCESSION	AX172854				
VERSION	AX172854.1	GI:14597903			
KEYWORDS					
SOURCE	Discosoma sp.				
ORGANISM	Discosoma sp.				
REFERENCE	1				
AUTHORS	Tsien,R.Y. and Gonzalez,J.E.				
TITLE	Detection of transmembrane potentials by optical methods				
JOURNAL	Patent: WO 0142211-A 6 14-JUN-2001;				
FEATURES	The Regents of the University of California (US)				
Source	location/Qualifiers				
	1..678				
	/organism="Discosoma sp."				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:86600"				

ORIGIN

Alignment Scores:					
Pred. No.:	1.4e-114	Length:	678		
Score:	1214.00	Matches:	225		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		

US-10-006-922A-12 (1-225) x AX172854 (1-678)

QY	1	MetArgSerSerLySAAnValIleLySGluPheMetArgPheLySValArgMetGluGly	20
DB	1	ATGAGGCTCTCCAGAAGATGTTATCAAGAGTTCATGAGGTTTAAGTTTCGCATGAGAGA	60
QY	21	ThrValAAnGLyHISgluPheGluIleGluGlyGluGlyArgProTyrgluGly	40
DB	61	ACGGTCAATGGGACGAGTTTGAATAGAAGCGCAAGAGGGGAGGCCATACGAAGGC	120
QY	41	HISAnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTPAapIle	60
DB	121	CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrglySerLySValTyrrVallySHISProAlaApIlePro	80
DB	181	TTGTCAACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240
QY	81	AspTyrrLySlySLeuSerPheProGluGlyPheLySTrpGluArgValMetAAnPheGlu	100
DB	241	GACTATAAAAAGCTGTATTTCTCTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySpheIleTyrr	120



Db 301 GACGGTGGCGTCTACTGTAAACCAGGATTCCAGTTTGCAAGATGGCTGTTTCACTCTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCCCTCCGATGGAAGCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGACCGTTGTATCCTCGTATGGCGGTGTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGGTCATTACCTTAGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAAAGAGCCCTGTGCAGCTAACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAAACAAGCCACAACGAAGACTATCAATCGTTGAGCAGTATGAAGAAGAACCGAGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTCTCCTT 675

RESULT 3  
AX207715 678 bp DNA linear PAT 31-AUG-2001  
LOCUS AX207715  
DEFINITION Sequence 13 from Patent WO0157242.  
ACCESSION AX207715  
VERSION AX207715.1 GI:15422399  
KEYWORDS

SOURCE Discosoma sp.  
ORGANISM Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE 1  
AUTHORS Stack,J.H., Whitney,M., Cubitt,A.B. and Pollok,B.A.  
TITLE Methods of protein destabilization and uses thereof  
JOURNAL Patent: WO 0157242-A 13 09-AUG-2001,  
Aurora Biosciences Corporation (US)

FEATURES  
source 1..678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX207715 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAGAATGTTATCAAGAGATTCAAGGTTTAAGGTTGCATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCAAGAGTTTGAATAAGAGCGAAGAGAGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGAGCCTTGGCATTTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 181 TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAGCTGTTCATTCTCGTAAGGATTTAATGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCCAGTTTGCAAGATGGCTGTTTCACTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCCCTCCGATGGAACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGGTCATTACCTTAGTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAAAGAGCCCTGTGCAGCTAACAGGGTACTACTATGTTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAAACAAGCCACAACGAAGACTATCAATCGTTGAGCAGTATGAAGAAGAACCGAGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTCTCCTT 675

RESULT 4  
AX233581 678 bp DNA linear PAT 11-SEP-2001  
LOCUS AX233581  
DEFINITION Sequence 5 from Patent WO0162919.  
ACCESSION AX233581  
VERSION AX233581.1 GI:15593305  
KEYWORDS

SOURCE Discosoma sp.  
ORGANISM Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.

REFERENCE 1  
AUTHORS Nelson,D., Zamaira,E. and Tsien,R.  
TITLE Modified fluorescent proteins  
JOURNAL Patent: WO 0162919-A 5 30-AUG-2001;  
Aurora Biosciences Corporation (US)

FEATURES  
source 1..678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
/note="red"

ORIGIN

Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX233581 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAGAATGTTATCAAGAGATTCAAGGTTTAAGGTTGCATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40



Db 61 ACGGTCAATGGGACGAGTTTGAATAGAAGCGGAGAGGAGGGAGCCATACGAAGGC 120  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATAACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACTCGCCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAGCTGTCAATTTCCTGAAGGATTTAAATGGGAAAGGCTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTGGTACTGTAAACCAAGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysLysThr 140  
Db 361 AAGGTCAAGTTCATTGGCGGTGAATTCTTCCTCCGATGGACCTGTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIu 160  
Db 421 ATGGGCTGGGAAGCCAGCACCTGAGCGTTTGATCCTCGTATGGCGGTGTGAAGGAGAG 480  
QY 161 IleHisLysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCACTTACCTTAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAAAGAGCCTGTGCACTACAGGGTACTACTATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACACAGAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTCTCCTT 675

RESULT 5  
AX370404 678 bp DNA linear PAT 16-FEB-2002  
LOCUS AX370404  
DEFINITION Sequence 1 from Patent WO0196373.  
ACCESSION AX370404  
VERSION AX370404.1 GI:18857490  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Pradkov,A.F. and Tersikh,A.  
TITLE Fluorescent timer proteins and methods for their use  
JOURNAL Patent: WO 0196373-A 1 20-DEC-2001;  
Clontech Laboratories Inc. (US)  
FEATURES  
Source location/Qualifiers  
1. .678  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="variant of sequence from Discosoma sp."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.4e-114 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
GB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX370404 (1-678)  
QY 1 MetArgSerSerLysAsnValIleIysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGCCCTCCTCCAAGACGTCAATCAAGAGATTCTATCGCCTTCAAGGTGCCGATGAGGCC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTGGCTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCGCGACATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATCTTGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGTGACCGTGACCAAGGACTCTCTCCCTGCAGGACGGCTGCTTCACTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysLysThr 140  
Db 361 AAGGTGAAGTTCATGGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAAGAACCC 420  
QY 141 MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIu 160  
Db 421 ATGGGCTGGAGGCTTCCACCGAGCGCCTGTACCCCCGAGCGCGGTGCTGAAGGGAG 480  
QY 161 IleHisLysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGCACCTACCTGTTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCCAGAAGCCCGTGCAAGCTGCCCGGCTACTACTACGTGAGTCCAAGCTGGA 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGCCGCC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACTGTCTCTG 675

RESULT 6  
AX824725 678 bp DNA linear PAT 11-DEC-2003  
LOCUS AX824725  
DEFINITION Sequence 7 from Patent WO02068459.  
ACCESSION AX824725  
VERSION AX824725.1 GI:39750591  
KEYWORDS  
SOURCE .  
ORGANISM Discosoma sp.  
REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
TITLE Non aggregating fluorescent proteins and methods for using the same  
JOURNAL Patent: WO 02068459-A 7 06-SEP-2002;  
CoralIimorpharia; Discosomatidae; Discosoma.  
FEATURES  
Source location/Qualifiers  
1. .678  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"

ORIGIN



Alignment Scores:

Pred. No.:	1.4e-114	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AF168419 (1-678)

QY	1	MetArGSeSerLybAsnValIlleYsGluPhemetaArgPheLybValArgMetGluGly	20
DB	1	ATGCGCTCTCCAAAGAACGTTCATCAAGGAGTTCAATGCGCTTCAAGGTGCGCATGAGGCG	60
QY	21	ThrValAaNGlYHISgluPheGluIlleGluGlyGluGlyArgProTYrGluGly	40
DB	61	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC	120
QY	41	HisAsnThrValLybLeuLybValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCTTGCCTGGACATC	180
QY	61	LeuSerProGlnPheGlnTYrGlySerLybValTYrValAllySHisProAlaAspIlePro	80
DB	181	CTGTCCCCCAAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCACATCCCC	240
QY	81	AspTYrLybLybLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPheGlu	100
DB	241	GACTACAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYr	120
DB	301	GACGGCGCGGTGTGACCGTGACCCAGGACTCTCCCTGCAGAGCGGCTGCTTCACTAC	360
QY	121	LybValLybPheIlleGlyValAsnPheProSerAspGlyProValMetGlnLybLySThr	140
DB	361	AAGGTGAAGTTCATCGCGCGTGAACCTCCCTCCGACGGCCCCGTGATGCAAGAAAGACC	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLySGlyGlu	160
DB	421	ATGGGCTGGAGGCCCTCCACCGAGCGGCTGTACCCCGGACGGCGGTGTAAGGGCGAG	480
QY	161	IlleHisLybAlaLeuLybLeuLybAspGlyGlyHisTYrLeuValGluPheLybSerIle	180
DB	481	ATCCACAAGGCCCTGAAGCTGAAGAGCGGGCCACTACCTGCTGAGTTCAAGTCCATC	540
QY	181	TYrMetAlaLybLybProValGlnLeuProGlyTYrTYrTYrValAspSerLybLeuAsp	200
DB	541	TACATGGCCAGAAGCCCGTGCAGCTGCCGCTACTACTACGTGACTCCAAAGCTGGAC	600
QY	201	IlleThrSerHisAaNGluAspTYrThrIleValGluGlnTYrGluArgThrGluGlyArg	220
DB	601	ATCACCTCCCAACAAGAGACTACACCATCGTGAAGACAGTACGAGCGCACCGAGGGCGC	660
QY	221	HisHisLeuPheLeu 225	
DB	661	CACCACCTGTCTCTG 675	
RESULT 7			
AF168419		859 bp	mRNA linear INV 27-JUL-2001
LOCUS	AF168419		Discosoma sp. fluorescent protein FP583 mRNA, complete cds.
DEFINITION	AF168419		
ACCESSION	AF168419.2	GI:7105733	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL	Nat. Biotechnol. 17 (10), 969-973 (1999)
PUBMED	10504696
REFERENCE	2 (bases 1 to 859)
AUTHORS	Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.
TITLE	Submitted (14-JUL-1999) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia
JOURNAL	3 (bases 1 to 859)
REFERENCE	Matz,M.V., Fradkov,A.F., Labas,Y.A., Savitsky,A.P., Zaraisky,A.G., Markelov,M.L. and Lukyanov,S.A.
AUTHORS	Direct Submission
TITLE	Submitted (25-FEB-2000) Institute of Bioorganic Chemistry RAS, Mklukho-Maklaya 16/10, Moscow 117871, Russia
REMARK	Sequence update by submitter
COMMENT	On Feb 25, 2000 this sequence version replaced gi:6090866.
FEATURES	location/Qualifiers
source	1. .859
	/organism="Discosoma sp."
	/mol_type="mRNA"
	/db_xref="taxon:86600"
	54. .731
	/note="GFP-like protein; orange-red"
	/codon_start=1
	/product="fluorescent protein FP583"
	/protein_id="AAF03369.1"
	/db_xref="GI:6090867"
	/translation="MRSSKNVITKEFMRPKVRMEGTVNGHEFEIEGEGRPYEGHNTV
	KLKVTKGPLPFAMDILSPQFGSKYVYKRPADIPDYKLSFPEGFKMERVANNEDG
	GVTVTQDSSLQDGCFTYKVFIGNVPSDGPVQKTMGWEASTERLYPRDGLKGE
	IRKALKLKDGGHYLVFESKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYRTE
	GRHHLFL"

ORIGIN

Alignment Scores:	1.87e-114	Length:	859
Pred. No.:	1214.00	Matches:	225
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	2	Gaps:	0
DB:			
US-10-006-922A-12 (1-225) x AF168419 (1-859)			
QY	1	MetArGSeSerLybAsnValIlleYsGluPhemetaArgPheLybValArgMetGluGly	20
DB	54	ATGAGGCTTCCAAAGATGTATCAAGAGTTCAAGGTTTAAGTTGCGCATGGAAGGA	113
QY	21	ThrValAaNGlYHISgluPheGluIlleGluGlyGluGlyArgProTYrGluGly	40
DB	114	ACGGTCAATGGGCAAGAGTTGAATAAGAAGCGAAGAGAGAGGGGAGCCATACGAAGGC	173
QY	41	HisAsnThrValLybLeuLybValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	174	CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGAGCCTTGCCATTGCTTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTYrGlySerLybValTYrValLySHisProAlaAspIlePro	80
DB	234	TTGTCAACCAAACTTCAGTATGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA	293
QY	81	AspTYrLybLybLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPheGlu	100
DB	294	GACTATAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTYr	120
DB	354	GACGGTGGCGTCACTGTAAACCCAGAGATTCAGATTGCAAGATGGCTGTTTCATCTAC	413
QY	121	LybValLybPheIlleGlyValAsnPheProSerAspGlyProValMetGlnLybLySThr	140
DB	414	AAGGTCAAGTTCATTGGCGTGAACCTTCTCCGATGGAACCTGTATTGCAAAAGAAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLySGlyGlu	160



|||||  
Db 474 ATGGCTGGGAAGCCAGCACTGACGCTTTGTATCCTCGTGATGCCGTGTGAAGAGAG 533  
Qy 161 ILEHISLYSALALEULYSLEULYSASPGLYGLYHISLYRLEUVALGLUPHELYSERILE 180  
Db 534 ATTCAITRAGGCTCTGAAGCTGAAAGACGGGTGTCATTACCTAGTTGAATTCAAAGTATT 593  
Qy 181 TYMETALALYSLYSPROVALGlnLEUPROGLYTYR TYR TYR VALASPSELYSLEUASP 200  
Db 594 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 653  
Qy 201 ILETHRSEHISAANGLUASP TYR THR ILE VALGLUGLINTYRGUARGTARGLUYARG 220  
Db 654 ATAAACAAGCCACAACGAAGACTATCAATCGTTGAGCAGTATGAAGAACCAGAGGAGCG 713  
Qy 221 HISHISLEUPHELEU 225  
Db 714 CACCATCTGTCTCCTT 728  
RESULT 8  
AR636082  
LOCUS AR636082 859 bp DNA linear PAT 14-FEB-2005  
DEFINITION Sequence 11 from patent US 6852849.  
ACCESSION AR636082  
VERSION AR636082.1 GI:59795931  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 859)  
AUTHORS Tsien,R.Y. and Campbell,R.E.  
TITLE Non-oligomerizing tandem fluorescent proteins  
JOURNAL Patent: US 6852849-A 11 08-FEB-2005;  
The Regents of the University of California; Oakland, CA  
FEATURES  
source 1..859  
/organism="Unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.87e-114 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922A-12 (1-225) x AR636082 (1-859)  
Qy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
Db 54 ATGAGGTCCTTCCAAGATGTTATCAAGAGGTTCATGAGGTTTAAGGTTTCGCATGGAAGCA 113  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
Db 114 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGGAGGAGGCCATACGAAGGC 173  
Qy 41 HISAAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATAACCGTAAGCTTAAGGTAAACCAAGGGGAGACCTTGGCCATTGGGATATT 233  
Qy 61 LeuSerProGlnPheGlnTyrglySerLySValTyrrValLyHisProAlaAspIlePro 80  
Db 234 TTGTCAACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCGACATACCA 293  
Qy 81 AspTyrrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAAAGCTGTCAATTTCTGAAGGATTTAATGGGAAAGGGTCATGAACCTTGAA 353  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrr 120  
Db 354 GACGGTGGCGTGTACTGTAAACCAAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC 413

Qy 121 LYSVALLYSPHEILEGLYVALAsnPheProSerAspGlyProValMetGlnLySerThr 140  
Db 414 AAGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTATGCAAAAGAGACA 473  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLySGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGATGGCGGTGAAAGAGAG 533  
Qy 161 ILEHISLYSALALEULYSLEULYSASPGLYGLYHISLYRLEUVALGLUPHELYSERILE 180  
Db 534 ATTCATTAAGGCTCTGAAGCTGAAAGACGGGTGTCATTACCTAGTGAATTCAAAGTATT 593  
Qy 181 TYMETALALYSLYSPROVALGlnLEUPROGLYTYR TYR TYR VALASPSELYSLEUASP 200  
Db 594 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAT 653  
Qy 201 ILETHRSEHISAANGLUASP TYR THR ILE VALGLUGLINTYRGUARGTARGLUYARG 220  
Db 654 ATAAACAAGCCACAACGAAGACTATCAATCGTTGAGCAGTATGAAGAACCAGAGGAGCG 713  
Qy 221 HISHISLEUPHELEU 225  
Db 714 CACCATCTGTCTCCTT 728  
RESULT 9  
AR670157  
LOCUS AR670157 859 bp DNA linear PAT 13-JUN-2005  
DEFINITION Sequence 11 from patent US 6900304.  
ACCESSION AR670157  
VERSION AR670157.1 GI:67608896  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 859)  
AUTHORS Tsien,R.Y., Ting,A.Y. and Zhang,J.  
TITLE Emission ratiometric indicators of phosphorylation  
JOURNAL Patent: US 6900304-A 11 31-MAY-2005;  
The Regents of the University of California; Oakland, CA  
FEATURES  
source 1..859  
/organism="Unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.87e-114 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922A-12 (1-225) x AR670157 (1-859)  
Qy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
Db 54 ATGAGGTCCTTCCAAGATGTTATCAAGAGGTTCAAGAGGTTTAAGGTTTCGCATGGAAGCA 113  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
Db 114 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGGAGGAGGCCATACGAAGGC 173  
Qy 41 HISAAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATAACCGTAAGCTTAAGGTAAACCAAGGGGAGACCTTGGCCATTGGGATATT 233  
Qy 61 LeuSerProGlnPheGlnTyrglySerLySValTyrrValLyHisProAlaAspIlePro 80  
Db 234 TTGTCAACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCGACATACCA 293  
Qy 81 AspTyrrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100



|||||  
Db 294 GACTATAAAAGCTGTCATTTCTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA 353  
QY 101 AaPglYglYValValThrValThrglnAaPseSerSerLeuGlnAaPglYcysPheIleTyr 120  
Db 354 GACGGTGGCGTCGTACTGTACCCAGGATTCAGTTTGcAGGATGGCTGTTTCATCTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAaPglYProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACA 473  
QY 141 MetGlyTrpGluAlaSerThrgluArgLeuTyrProArgAaPglYValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAGAGAG 533  
QY 161 lLeHleLysAlaLeuLysLeuLysAaPglYglYHlsTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAaPseSerLysLeuAaP 200  
Db 594 TACATGGCAAGAAGCCTGTGCAAGCTACCAAGGCTACTACTATGTTGACTCCAAACTGGAT 653  
QY 201 lLeThSerHlsAaNgLuAaPlyTyrThrlleValGluGlnTyrGluArgThrgluGlyArg 220  
Db 654 ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCGAAGGAGCGC 713  
QY 221 HlsHlsLeuPheLeu 225  
Db 714 CACCATCTGTTCCCTT 728

RESULT 10  
AX463698 859 bp DNA linear PAT 15-JUL-2002  
LOCUS AX463698 Sequence 12 from Patent WO0248338.  
DEFINITION AX463698  
ACCESSION AX463698  
VERSION AX463698.1 GI:21886457  
KEYWORDS  
SOURCE  
ORGANISM  
Discosoma sp.  
Discosoma sp.  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE  
AUTHORS  
TITLE  
Lichtenberg-Frat,H.  
1  
Yeast strain for testing the geno- and cytotoxicity of complex  
environmental contamination  
JOURNAL  
Patent: WO 0248338-A 12 20-JUN-2002;  
Lichtenberg-Frat, Hella (DE)  
FEATURES  
source  
1. 859  
/organism="Discosoma sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:86600"  
54..731  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD42147.1"  
/db\_xref="GI:21886458"  
/translation="MRSSKNVLIKEMRPFKVRMEGTVNGHEFEIIEGEGEGRPYEGHNTV  
KLKVTKGGLPFAWDILSPQFOYGSKYVVKHPADIPDYKKLSPEEGFKMERVMNFEDG  
GVVTVQDSSLQDGCFTYKVKFISGVNFPDSDGPMQKKTGMWEASLERLYPRDGLKGE  
IHKALKKLDGGHYLVFEKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTE  
GRHMLPL"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.87e-114 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX463698 (1-859)  
QY 1 MetArgSerSerLysAaValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGGTCCTTCCAAGATGTTATCAAGAGTTCATGAGTCTTAAGTTCGATGGAAGA 113  
QY 21 ThrValAaNgLyHlsGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCACGAGTTGAAATGAAGCGGAGAGAGAGGAGGCCATACGAAGGC 173  
QY 41 HlsAaThrValLysLeuLysValThrlYsgLylYProLeuProPheAlaTrpAaPle 60  
Db 174 CACATAACCGTAAAGCTTAAGGTAACCAAAGGGGAGCCTTTGCCATTGCTTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHlsProAlaAaPlePro 80  
Db 234 TTGTCAACCACAATTTCAGTATGGAACGAAGGTATATGTCAAGCACCTGCGACATACCA 293  
QY 81 AaPlyTrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAaPheGlu 100  
Db 294 GACTATAAAAGCTGTGTCATTTCTGAAGATTTAATGGGAAAGGTCATGAACCTTGAA 353  
QY 101 AaPglYglYValValThrValThrglnAaPseSerSerLeuGlnAaPglYcysPheIleTyr 120  
Db 354 GACGGTGGCGTCGTACTGTACCCAGGATTCAGTTTGcAGGATGGCTGTTTCATCTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAaPglYProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACA 473  
QY 141 MetGlyTrpGluAlaSerThrgluArgLeuTyrProArgAaPglYValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAAGAGAG 533  
QY 161 lLeHleLysAlaLeuLysLeuLysAaPglYglYHlsTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAaPseSerLysLeuAaP 200  
Db 594 TACATGGCAAGAAGCCTGTGCAAGCTACCAAGGCTACTACTATGTTGACTCCAAACTGGAT 653  
QY 201 lLeThSerHlsAaNgLuAaPlyTyrThrlleValGluGlnTyrGluArgThrgluGlyArg 220  
Db 654 ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCGAAGGAGCGC 713  
QY 221 HlsHlsLeuPheLeu 225  
Db 714 CACCATCTGTTCCCTT 728

RESULT 11  
AX233584 681 bp DNA linear PAT 11-SEP-2001  
LOCUS AX233584 Sequence 8 from Patent WO0162919.  
DEFINITION AX233584  
ACCESSION AX233584  
VERSION AX233584.1 GI:15593307  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Nelson,D., Zamaira,E. and Tsielen,R.  
Modified fluorescent proteins  
Patent: WO 0162919-A 8 30-AUG-2001;  
Aurora Biosciences Corporation (US)  
FEATURES  
source  
1. 681  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"  
1..681  
/note="unnamed protein product"



/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC69734.1"  
/db\_xref="GI:15593308"  
/translation="MVRSSKNVKEFMRFKVRMEGTVNGHEFEIEGEGRPEYEGHNT  
VKLKVTKGGLPFAMDLSPQYQYKVVKHPADIPTYKLSFPEGFKWERVMNFED  
GGVTVTQDSSLQDGCIFYKVKFIGNFPSPDGPVMQKKTMGWEASTERLYPRDGLKG  
EIHKALKDKDGHYLVFESKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIWEQYERT  
EGRHHLFL"

ORIGIN

Alignment Scores:

Pred. No.:	3.61e-114	Length:	681
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AX233584 (1-681)

QY	1	MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
DB	4	GTGAGGAGCAGCAAGACGTGATCAAGAGTTTCATGAGGTTCAAGGTGCGCATGAGGGC	63
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	64	ACCGTGAAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGCCCTACGAGGGC	123
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
DB	124	CACAACACCGTGAGCTTAAGGTGACCAAGGGCGGCCCTCGCCCTTGCGTGGACATC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
DB	184	CTGAGCCCCCAGTTCAGTACGGGCAAGGTTGTAAGTGAAGCAACCCCGCGACATCCCC	243
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	244	GACTACAAGAGCTGAGCTTCCCCGAGGGCTTCAAGTGGAGAGGGGTGATGAACCTTCGAG	303
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	304	GACGGCGGCGTGTGACCGTGACCCAGACAGACAGCAGCCTGCAGAGCGGCTGCTTCACTTAC	363
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
DB	364	AAGGTGAAGTTCATCGGCGGTGAACCTTCCCCAGCGACGGCCCGTGATGCAGAAAGAACCC	423
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySLeuGlu	160
DB	424	ATGGGCTGGGAGGCGCTCCACCGACGGCCTGTACCCCCCGGACGGCGTGTGAAGGGCGAG	483
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180
DB	484	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGCTGGTGAAGTTCAAGTCCATC	543
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	544	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGACTCCAAGCTGGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	604	ATCACCAAGCCACAACGAGGACTACACCATCGTGAGCAGTACGAGAGGACCGAGGGCGAGG	663
QY	221	HisHisLeuPheLeu	225
DB	664	CACCACTGTCTCTG	678

RESULT 12  
AX233627  
LOCUS AX233627 713 bp DNA linear PAT 11-SEP-2001  
DEFINITION Sequence 51 from Patent WO0162919.

ACCESSION AX233627  
VERSION AX233627.1 GI:15593330  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
Source  
1. 713  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Mutant Anthozoan red fluorescent protein"

ORIGIN

Alignment Scores:

Pred. No.:	3.81e-114	Length:	713
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AX233627 (1-713)

QY	1	MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly	20
DB	22	GTGAGGAGCAGCAAGACGTGATCAAGAGTTTCATGAGGTTCAAGGTGCGCATGAGGGC	81
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	82	ACCGTGAAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGCCCTACGAGGGC	141
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
DB	142	CACAACACCGTGAGCTTAAGGTGACCAAGGGCGGCCCTCGCCCTTGCGTGGACATC	201
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
DB	202	CTGAGCCCCCAGTTCAGTACGGGCAAGGTTGTAAGTGAAGCAACCCCGCGACATCCCC	261
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	262	GACTACAAGAGCTGAGCTTCCCCGAGGGCTTCAAGTGGAGAGGGGTGATGAACCTTCGAG	321
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	322	GACGGCGGCGTGTGACCGTGACCCAGACAGACAGCAGCCTGCAGAGCGGCTGCTCATCTAC	381
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
DB	382	AAGGTGAAGTTCATCGGCGGTGAACCTTCCCCAGCGACGGCCCGTGATGCAGAAAGAACCC	441
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySLeuGlu	160
DB	442	ATGGGCTGGGAGGCTTCCACCGAGCGCTGTACCCCCCGGACGGCGTGTGAAGGGCGAG	501
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180
DB	502	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGCTGAGTTCAAGTCCATC	561
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	562	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGACTCCAAGCTGGAC	621
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	622	ATCACCAAGCCACAACGAGGACTACACCATCGTGAGCAGTACGAGAGGACCGAGGGCGAGG	681



Qy	221	HisHisLeuPheLeu	225	
Db	682	CACCACCTGTTCTG	696	
RESULT 13				
LOCUS	AR527331		723 bp	DNA
DEFINITION	Sequence 1 from patent US 6723537.			
ACCESSION	AR527331			
VERSION	AR527331.1	GI:53914309		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unpublished.			
AUTHORS	1 (bases 1 to 723)			
TITLE	Peelle, B.			
JOURNAL	Directed evolution of protein in mammalian cells			
	Patent: US 6723537-A 1 20-APR-2004;			
	Rigel Pharmaceuticals, Incorporated and Becton, Dickinson and			
	Company; South San Francisco, CA			
FEATURES				
source	1..723			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Alignment Scores:				
Pred. No.:	3.88e-114	Length:	723	
Score:	1210.00	Matches:	224	
Percent Similarity:	100.00%	Conservative:	1	
Best Local Similarity:	99.56%	Mismatches:	0	
Query Match:	99.67%	Indels:	0	
DB:	6	Gaps:	0	
US-10-006-922A-12 (1-225) x AR527331 (1-723)				
Qy	1	MetArgSerSerLybAbnValIleLybGluPheMetArgPheLybValArgMetGluLy	20	
Db	4	GTGCGCTCTCCAGAACGTCATCAAGAGTTTCATGCGCTTCAAGTGCATGAGGGC	63	
Qy	21	ThrValABNGlyHISGluPheGluIleGluGlyGluGlyGlyArgProTyrgLgLy	40	
Db	64	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTTACGAGGC	123	
Qy	41	HisAbnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle	60	
Db	124	CACAACACCGTGAAAGCTGAAGGTACCAAGGGGGCCCCCTTGCCTTGGGACATC	183	
Qy	61	LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybHisProAlaAspIlePro	80	
Db	184	CTGTCCCCCAGTTCAGTACGGCTTCCAAAGGTGACGTAGAACACCCCGGCACATCCCC	243	
Qy	81	AspTyrlLybLybLeuSerPheProGluGlyPheLybSTrgLuarGValMetAbnPheGlu	100	
Db	244	GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAG	303	
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTy	120	
Db	304	GACGGCGCGCGTGTGACCGCTGACCCAGACTCTCCCTGCAAGACGGCTGCTCATCTAC	363	
Qy	121	LybValLybPheIleGlyValAbnPheProSerAspGlyProValMetGlnLybLybThr	140	
Db	364	AAAGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTAATGACAAAGAAC	423	
Qy	141	MetGlyTrpGluAlaSerThrgLuarGLeuTyrgProArgAspGlyValLeuLybGlyGlu	160	
Db	424	ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCAGCGGCGTGTGAAGGGCGAG	483	
Qy	161	IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrlLeuValGluPheLybSerIle	180	
Db	484	ATCCACAAGCGCTGAAGCTGAAGACGGGGCCACTACCTGTGAGTTCAAGATATC	543	
Qy	181	TyrMetAlaLybLybProValGlnLeuProGlyTyrgTyrgTyrgValAspSerLybLeuAsp	200	

Db	544	TACATGGCCCAAGAGCCCGTCGACTGCCCGCTACTACTACTGACTCCAAGCTGGAC	603
Oy	201	IleThSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluYarg	220
Db	604	ATCACCCTCCCAACAAGAGACTACACCAATCGTGAGCAGTAGCAGCGCACCGGC	663
Oy	221	HISHisLeuPheLeu	225
Db	664	CACCACCTGTTCCTG	678
 RESULT 14			
LOCUS	CQ882115	2721 bp	DNA linear PAT 11-OCT-2004
DEFINITION	Sequence 1 from Patent WO2004083445.		
ACCESSION	CQ882115		
VERSION	CQ882115.1	GI:54034825	
KEYWORDS	synthetic construct synthetic construct other sequences; artificial sequences.		
SOURCE			
ORGANISM			
REFERENCE	1 Chavancy,G., Couble,P., Durand,B., Grenier,A.M., Horard,B., Julien,E., Mauchamp,B., Nony,P., Prudhomme,J.C. and Royer,C. Nucleic acid controlling the expression of a useful polypeptide in the posterior silk glands of a lepidoptera and application thereof Patent: WO 2004083445-A 1 30-SEP-2004; Centre National De la Recherche Scientifique-CNRS (FR); UNIVERSITE CLAUDE BERNARD - LYON 1 (FR); Institut National de la Recherche Agronomique (INRA) (FR)		
JOURNAL	Location/Qualifiers		
FEATURES	1..2721		
source	/organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /note="Description de la s quence artificielle : s quence de fusion"		

ORIGIN	
Alignment Scores:	
Pred. No.:	1.95e-113
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	6
	Gaps: 0
US-10-006-922A-12 (1-225) x CQ882115 (1-2721)	
QY	1 MetArgSerSerLybAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
DB	2044 GTGCGCTCTCCAGAACGTCAATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 2103
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB	2104 ACCGTGAACGCCCAAGAGTTGAGATCGAGGCGAGGGCGAGGGCCGCTTACGAGGGC 2163
QY	41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
DB	2164 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCTGCGCTTGGCTGGACATC 2223
QY	61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
DB	2224 CTGTCCCCCAGTTCCAGTAACGGCTCCAAAGGTATCGTGAAGCAACCCGCCGACATCCCC 2283
QY	81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100
DB	2284 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAG 2343
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB	2344 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGCAGGACGCGTGTCAATCTAC 2403



QY	121	LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr	140
Db	2404	AAGGTGAAGTTCATCGGCGGTGAACCTTCCCTCCGACGGCCCGGTATGCAGAGAAGACC	2463
QY	141	MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	2464	ATGGGCTGGAGGCGCTCCACCAGCGCGCTGTACCCCGGACGGCGGTGCTGAAGGGCGAG	2523
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisIstYrLeuValGluPheLysSerIle	180
Db	2524	ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACCTACTGTGGAGTTCAGTTCATC	2583
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	2584	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGACTCCAAGCTGCAC	2643
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	2644	ATCACCTCCCAACACGAGACTACACCATCTGTGAGCACTACGAGCGCACCGAGGGCCGC	2703
QY	221	HisHisLeuPheLeu	225
Db	2704	CACCACCTGTTCCTG	2718

RESULT 15			
CQ882117			
LOCUS	CQ882117	2772 bp	DNA
DEFINITION	Sequence 3 from Patent WO2004083445.	linear	PAT 11-OCT-2004

VERSION	Q8082117	GI:54034827
KEYWORDS	Q882117.1	
SOURCE	synthetic construct	
ORGANISM	synthetic construct	
	other sequences; artificial sequences.	

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Chavancy, G., Couble, P., Durand, B., Grenier, A.M., Horard, B., Julien, E., Mauchamp, B., Nony, P., Prudhomme, J.C. and Royer, C.	Nucleic acid controlling the expression of a useful polypeptide in the posterior silk glands of a lepidoptera and application thereof	Patent: WO 2004083445-A 3 30-SEP-2004;
	Centre National De La Recherche Scientifique-CNRS (FR) ; UNIVERSITE CLAUDE BERNARD - LYON 1 (FR) ; Institut National de la Recherche Agronomique (INRA) (FR)		

```

FEATURES
    source
        location/Qualifiers
            1. .2772
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Description de la s quence artificielle : s quence
de fusion"

```

ORIGIN	
Alignment Scores:	
Pred. No.:	2e-113
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	6
Length:	2772
Matches:	224
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x CQ882117 (1-2772)

Qy	1	MeLArgSerSerLysAsnValIleLysGluPhenMetArgPheLysValArgMetGluGly	20
	:::		
Db	2095	GTGCGCTCTCCAGAACGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG	2154
Qy	21	ThrValAsnGlyHseGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	2155	ACCGTGACCGGCCACGAGTTGAGATCGAGGGCGGAGGGCGGCCCTTACGAGGGC	2214
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	2215	CACACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTTGCCCTTGAGACATC	2274

QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	2275	CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCCCGACATCCCC	2334
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	2335	GACTACAAGAACTGTCTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCGAG	2394
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	2395	GACGGCGCGTGGTGAACCGTGACCAGGACTCTCCCTGCAGAGCAGCGCTGCTTCATCTAC	2454
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr	140
Db	2455	AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAATGCAGAAAGAACC	2514
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	2515	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCCGACGCGCGTGTGAAGGGCGAG	2574
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGlnPheLysSerIle	180
Db	2575	ATCCACAAGGCCCTGAAGCTGAAGAGCAGGGGCCACTACCTGCTGAGTCAAGTCCATC	2634
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	2635	TACATGGCCCAAGAGCCCGTGCACTGCCCGGCTACTACTACGTGACTCCAAGCTGAC	2694
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlnGlyArg	220
Db	2695	ATCACCTCCCAACAAGGACTACACCATCTGTGAGACAGTACGAGCGCACCGAGGGCCGC	2754
QY	221	HisHisLeuPheLeu	225
Db	2755	CACCACCTGTCTCTG	2769

RESULT 16	AX463702	LOCUS	AX463702	4692 bp	DNA	linear	PAT 15-JUL-2002
DEFINITION	Sequence 16 from Patent WO0248338.						
ACCESSION	AX463702						
VERSION	AX463702.1	GI:21886461					
KEYWORDS	.						
SOURCE	synthetic construct						
ORGANISM	synthetic construct						
	other sequences; artificial sequences.						

REFERENCE AUTHORS TITLE JOURNAL	1
Lichtenberg-Frat, H. Yeast strain for testing the geno- and cytotoxicity of complex environmental contamination Patent: WO 0248338-A 16 20-JUN-2002;	

FEATURES	Location/Qualifiers
source	1..4692
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="Vektor pBSRed1-N1"

# ORIGIN

Alignment Scores:	
Pred. No.:	3_79e-113
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	99.67%
DB:	6
Gaps:	0
Mismatches:	0
Conservative:	1
Matches:	224
Length:	4692

US-10-006-922A-12 (1-225) x AX463702 (1-4692)

QY 1 MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGly 20  
:::



Db	682	GTGCGCTCCTCCAGAACGTCATCAAGAGTTCATGCGCTTCAAGTGCGCATGAGGCG	741
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	742	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC	801
Qy	41	HisAsnThrValLysLeuValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	802	CACACACCGGTGAAGCTGAAGTGACCAAGGGCGGCCCTTCCTCGCGACATC	861
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	862	CTGTCCCCCACTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC	921
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	922	GACTACAAAGAAAGTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG	981
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	982	GACGGCGCGTGGTGACCGTGACCCAGGACTCTCTCCCTGCAGAGCGGCTTTCATCTAC	1041
Qy	121	LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr	140
Db	1042	AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC	1101
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	1102	ATGGGCTGGGAGGCTCTCCACCGAGCGCTGTATCCCCCGGACGGCGTGCTGAAGGGCGAG	1161
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	1162	ATCCACAGGCGCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGAAGTTCAAGTCCATC	1221
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	1222	TACATGGCCAGAAAGCCCGTGCAAGTCCCGGCTACTACTACGTGAGTCCAAAGCTGGAC	1281
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	1282	ATCACCTCCCAACAAGAGACTACACCATCGTGGAGCAGTACGAGCGCACCGAGGGCGGC	1341
Qy	221	HisHisLeuPheLeu 225	
Db	1342	CACCACTGTTCCTG 1356	
RESULT 17			
AX823860		6893 bp	DNA
LOCUS	AX823860		6893 bp
DEFINITION	Sequence 8 from Patent WO03070931.		
ACCESSION	AX823860		
VERSION	AX823860.1	GI:39750176	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	1		
AUTHORS	Baum,C., Will,E., Oebertag,W., Klump,H. and Schiedlmeier,B.		
TITLE	Methods for conducting site-specific dna recombination		
JOURNAL	Patent: WO 03070931-A 8 28-AUG-2003;		
FEATURES			
source			
	1. .6893		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Beschreibung der kuenstlichen Sequenz: provirale		
	plasmid-DNA; retrovirales Cre-Reporterkonstrukt.#SFr#"		
	1. .435		
	/note="plasmid-Rueckgrat (pUC) "		
	436. .996		
	/note="MPSV 5'-LTR (Delta1-31) "		
	997. .1014		
	primer_bind		

5'UTR	/note="PBS (primer binding site)"
misc_feature	1015. .1560
misc_feature	1561. .1599
misc_feature	/note="#loxP1#-Sequenz"
misc_feature	1600. .2289
misc_feature	/note="Derecl-Gen aus Discosoma sp. "
misc_feature	2290. .2343
misc_feature	/note="#loxP2#-Sequenz"
misc_feature	2347. .3150
misc_feature	/note="egFP(tag)-Gen aus Aequorea victoria"
misc_feature	3164. .4019
3'UTR	/note="#wPRE#"
misc_feature	4020. .4080
misc_feature	4081. .4641
misc_feature	/note="3'-LTR"
misc_feature	4642. .6893
misc_feature	/note="plasmid-Rueckgrat (pUC) "

ORIGIN

Alignment Scores:

Pred. No.:	6.06e-113	Length:	6893
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AX823860 (1-6893)

Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1603	GTGCGCTCTCCAAAGACGTATCAAGAGTTATGCGCTTCAAGTGCGCATGAGGCGC	1662
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	1663	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGAGGGCGGCCCTACGAGGCGC	1722
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	1723	CACACACCGGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCTTCCCTGGACATC	1782
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	1783	CTGTCCCCCACTTCCAGTACGGCTCCCAAGGTGTACGTGAAGCACCCCGACATCCCC	1842
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	1843	GACTACAAAGAACTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG	1902
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	1903	GACGGCGCGTGGTGACCGTGACCCAGGACTCTCCTCTGAGGACGGCTTTCATCTAC	1962
Qy	121	LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr	140
Db	1963	AAGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC	2022
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	2023	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGGACGGCGGTGTGAAGGGCGAG	2082
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	2083	ATCCACAAAGGCTTGAAGCTGAAGGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC	2142
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	2143	TACATGGCCAAAGACCGGTGACGCTGCCGCTACTACTACGTGACTTCCAAGCTGGAC	2202
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	2203	ATCACCTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCGGC	2262



Db	2263	CACCACCTGTTCTG	2277
LOCUS	AY569779	8811 bp	DNA
DEFINITION	Cloning vector pTCR, complete sequence.		
ACCESSION	AY569779		
VERSION	AY569779.1	GI:45861232	
KEYWORDS			
SOURCE	Cloning vector pTCR		
ORGANISM	Cloning vector pTCR		
REFERENCE	1 (bases 1 to 8811)		
AUTHORS	Wu, G., Wang, Y. and Zhu, Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-MAR-2004) Fish Genetics, Institute of Hydrobiology, Chinese Academy of Science, 7# South Road of East Lake, Wuhan, Hubei 430072, People's Republic of China		
FEATURES	Location/Qualifiers		
source	1..8811		
	/organism="Cloning vector pTCR"		
	/mol_type="other DNA"		
	/db_xref="taxon:268733"		
	419..452		
	/note="Loxp RE"		
misc_feature	complement(505..555)		
terminator	/note="SV40 polyA"		
CDS	complement(707..1387)		
	/codon_start=1		
	/transl_table=1		
	/product="RFP"		
	/protein_id="AAS78495.1"		
	/db_xref="GI:45861233"		
	/translation="MVRSSKNVKEFMRFKVRMEGTNGHEFEIEGEGRPYECHNTVKLKTKGGLPFPAMDILSPQFGSKYVKHPADIPDYKLSFPEGFKMERVMNPFEGVTVTQDSSLQDGCFFIKVKFIVNFPSPDGPVMQKTMGEASTERLYPRDGLKGEIHKALKLKDGGHYLVFESKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTEGRHILFL"		
promoter	complement(1447..2036)		
	/note="CMV promoter"		
CDS	complement(2387..2761)		
	/codon_start=1		
	/transl_table=1		
	/product="zeocin"		
	/protein_id="AAS78496.1"		
	/db_xref="GI:45861234"		
	/translation="MAKLTSAVPVLTARDVAGAVEFWTDRLGFSRDFVEDDFAGVRDDVTLFISAVQDQVPDNTLAWVWVRGLDELVAEWSVSTNFRDASGPAMTEIGEOPWGRFALRDPAGNCVHFVAEEQD"		
terminator	complement(2860..2992)		
	/note="SV40 polyA"		
CDS	complement(3066..4097)		
	/codon_start=1		
	/transl_table=1		
	/product="Cre"		
	/protein_id="AAS78497.1"		
	/db_xref="GI:45861235"		
	/translation="MANLLTVHQNLPAIPVDATSDVRKNLMDMFRDQAFSEHTWKMILSVCSWAAWCKLNNRKWFPAEPEDVRYLLYLQARGLAVKTIQHLGOLMLHRRSGLPRPSDNASVSLVMRIRIKENVDAGERAKQALAFERTDFDQVRSIMENSRCODIRNLAPLGIAYNTLLRIABIRIRIKVDISRTDGGRLIHIGRTKLVSTAGVEKALSIGVTKLVERMTSVSGVADDPNNYLFQVRKNGVAAPSATSQLSTRALBEGIFEAITHRLIYGAKDMSGORYLAWSGHSARVGAARDMARAGVSIPEIMQAGWTNNIIVNNYIRNLDSETGAMVRLLEDGD"		
misc_feature	complement(4109..4683)		
	/note="internal ribosome entry sequence from encephalomyocarditis virus"		
promoter	complement(4745..4762)		
	/note="T7 promoter"		
CDS	complement(5151..6233)		
	/codon_start=1		

```

/misc_feature
rep_origin
CDS

/transl_table=11
/product="laci"
/protein_id="AAS78498.1"
/db_xref="GI:45861236"
/translation="MKPVTLYDVAEYAGVSQYTVSRVNVQASHVSAKTRKEVEAMAE
LNYIPNRVAQOLAGKOSLIGVATSSIALHAPSQIVAAIKSRADQIGASVWSMVER
GVEACKAAVHNLAQRVSGELINYPPLDDQDAI AVEAECTNVPALFLDVSQTPINSII
FSHEDGTRLGVEHLVALGHQOIALPLDSSVSARLRLAGWHKYLTRNQIQPIAEREG
DWSAMSGFQQTQMQLNEGIVPTAMLVANDQMALGAMRAITSEGLRVGADISVGYYDDT
EDSSCYIPSTTIKQDFRLQGTSDVRLQLQSGQAVKGNQLPEVSLVKRKTTLAPNT
QTASPRALADSLMQLARQVSRLESQ"
6520..6553
/note="loxP LE"
6992
/note="Col E1"
complement(7751..8611)
/codon_start=1
/transl_table=11
/product="Amp"
/protein_id="AAS78499.1"
/db_xref="GI:45861237"
/translation="MSIQHFRVALIPFFAFCFLPVFAHPETLVKVDADQLGARVGY
IELDIANGKILLESFRPEERPFMMSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAATMSDNTAANLLTTIGPKELTAFIAHMGDHTVRL
DRWEPELNEAIPNDERDTWVAMATTIRKLLTGSELLTASRQQLIDWMEADKVAQPL
LRSAIPAGMFIADKSGAGSGSRGI IALGPDGKPSRIIVITYTGSQATMDERRNRI A
EIGASLIKHW"

```

Alignment Scores:			
Pred. No.:	8.18e-113	Length:	8811
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922A-12 (1-225) x AYS69779 (1-8811)			
QY	1 MetArgSerSerLyAsnValIleYsGluPheMetArgPheLySValArgMetGluGly	20	
DB	1384 GTGCGTCTCCCAAGACGTCAAGAGATTATGCGCTTCAAGGTGCGCATGAGGGC	1325	
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40	
DB	1324 ACCGTGAACGGCCACGAGTTCGAGTCAGAGGGCGAGGGCGGGCCCTACGAGGGC	1265	
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheIaIaTrpAspIle	60	
DB	1264 CACACACCGTGGAAGCTGAAGGTGACCAAGGGCGGGCCCTGCGCTTGGGACATC	1205	
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProIaAspIlePro	80	
DB	1204 CTGTCCCCCAGTTCAGTACGGCTCCAGGTGTACGTGAAGCACCCCGGCACATCCCC	1145	
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100	
DB	1144 GACTACAGAAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCCAG	1085	
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120	
DB	1084 GACGGCGCGTGTGACCGTGACCCAGACTCTCTCCCTGCAGACGCGTGCATCTAC	1025	
QY	121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140	
DB	1024 AAGGTGAAGTTCATCGCGGTGAATTCTCCCTCCGACGGCCCCGTAAATGCAGAGAAGACC	965	
QY	141 MetGlyTrpGluIaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160	
DB	964 ATGGCTGGGAGGCTTCAACCGACGCGCTGTACCCCGGACGGCGGTGTGAAGGGCGAG	905	
QY	161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180	



Db 904 ATCCACAAGCCCTGAAGCTGAAGGACGGCGGCCACTACTGCTGGAGTTCAAGTCCATC 845

Qy 181 TYNMATAIAlYbLySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 844 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTAGTGAAGTCCAAAGCTGGAC 785

Qy 201 ILeThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220

Db 784 ATCACTCCCAACAAGAGACTACCATCGTGGAGCAGTACGACGCGCACCGAGGGCCGC 725

Qy 221 HisHisLeuPheLeu 225

Db 724 CACCACTGTTCCTG 710

RESULT 19

AX663075 LOCUS AX663075 9320 bp DNA linear PAT 24-MAR-2003

DEFINITION Sequence 20 from Patent WO02070740.

ACCESSION AX663075

VERSION AX663075.1 GI:29169369

KEYWORDS

SOURCE Synthetic construct

ORGANISM Synthetic construct

REFERENCE 1 Wiesmuller, L.

AUTHORS Test system for determining gene toxicities

TITLE Patent: WO 02070740-A 20 12-SEP-2002;

JOURNAL Wiesmuller, Lisa (DE)

FEATURES

source Location/Qualifiers

1. .9320

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="Plasmid p5-Puro-CMV- (N'-EGFP) -CMV-Red- (EGFP-EU) "

1. .1592

/note="Retroviraler Vektor p5NM"

1617. .2216

/note="unnamed protein product; Puromycin-Resistenzgen aus

Pretroon (Clontech, Palo Alto, CA, USA) "

/codon\_start=1

/transl\_table=11

/protein\_id="CAD80174.1"

/db\_xref="GI:29169370"

/translation="MTBYKPTVRLATRDVPRAVRTLAAAPADYPATRHVTDPDRHIE

RVTLEQLFLTRVGLDIGKWWADGAVAVWTTPESEVAGAVFAIGPRMAELSGSR

LAAQOMEGLLAPHRPKPEPAFLATVGVSPDHQKGLGSAVLLPGVEAERAGVPAFL

ETSAPRNLPEYERLIGFTVADVECPDRATWCMTRKPGA"

2267. .2848

/note="CMV-Promotor aus pEGFP-N1 (Clontech, Palo Alto, CA,

USA) "

2906. .>3349

/note="unnamed protein product; N'-EGFP, abgeleitet von

EGFP aus pEGFP-N1 (Clontech, Palo Alto, CA, USA) "

/codon\_start=1

/transl\_table=11

/protein\_id="CAD80175.1"

/db\_xref="GI:29169371"

/translation="MVSKEELFTGVNPILVELDGVNGHKFSVSGEGEDATYGKLT

LKFICTTGKLPVPMPTLVTTLTLYGVCFSRYPDHMKQHDFKSAPEGYVQERTIFK

DDGNYKTRAIEVKFEGDTLVNRIELKGIIDFKEDGNIIGHKLEVYNR"

3374. .3392

/note="Retroviraler Vektor p5NM"

3411. .3992

/note="CMV-Promotor aus pBSRed1-N1 (Clontech, Palo Alto,

CA, USA) "

4038. .4718

/note="unnamed protein product; Red aus pBSRed1-N1

(Clontech, Palo Alto, CA, USA) "

/codon\_start=1

/transl\_table=11

/protein\_id="CAD80176.1"

/db\_xref="GI:29169372"

CDS

misc\_feature

promoter

/translation="MYRSSKNVIKEFMRFKVRMEGTUNGHEFEIEGEGGRPYEGHNT

VKLKVTKGGLPFAWDILSPQFYGSKYVYKHPADIPDYKLSFPEGFKRVRVNFED

GGVTVTQDSSLQDCEIYKVKFIQVNFPSDGPWQKTMGWEASTRLYPRDGLKG

BIHKALKDKDGHYLVFEKSIYMAKRPVQLPGYYVDSKLDITSHNEDYTIIVEQYERT

EGRHHLFL"

4766. .5509

/note="EGFP-EU, abgeleitet von EGFP aus pEGFP-N1

(Clontech, Palo Alto, CA, USA) "

5527. .9320

/note="Retroviraler Vektor p5NM"

ORIGIN

Alignment Scores:

Pred. No.: 8.76e-113 Length: 9320

Score: 1210.00 Matches: 224

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 99.56% Mismatches: 0

Query Match: 99.67% Indels: 0

DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AX663075 (1-9320)

Qy 1 MetArgSerSerLySAsnValIleLySGLuPheMetArgPheLySValArgMetGluGly 20

Db 4041 GTGCGCTCTCCAAAGACGTCAATCAAGAGAGTTTCATGCGCTTCAAGTGGCATGAGAGGC 4100

Qy 21 ThrValAsnGlyHisGluPheGlnIleGluGlyGluGlyGluGlyArgProTyrGluGly 40

Db 4101 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGC 4160

Qy 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTyrAspIle 60

Db 4161 CACAACACGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTGCGTGGACATC 4220

Qy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80

Db 4221 CTGTCCCCCAGTTCAGTACGGCTCAAGGTGACGTGAAGCACCCCGACATCCCC 4280

Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100

Db 4281 GACTACAAGAAGCTGTCTTCCCCGAGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 4340

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 4341 GACGGCGGCTGTGACCGTGACCCAGGACTCTCTCCGCAAGACGGCTTCACTAC 4400

Qy 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140

Db 4401 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCGTAAATGCAAGAAAGACC 4460

Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160

Db 4461 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGGTGAAAGGCGAG 4520

Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrIleValGluPheLySserIle 180

Db 4521 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTACTGAGTCAAGTCCATC 4580

Qy 181 TyrMetAlaLySbProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 4581 TACATGGCCAAAGAGCCCGTGACGCTGCCGCTACTACTACGTGACTCCAAGCTGGAC 4640

Qy 201 ILeThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220

Db 4641 ATCACTCCCAACAAGAGACTACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 4700

Qy 221 HisHisLeuPheLeu 225

Db 4701 CACCACTGTTCCTG 4715

RESULT 20

AY569780/c LOCUS AY569780 12404 bp DNA circular SYN 05-APR-2004



FEATURES	source
DEFINITION	Cloning vector pTCR-Auto, complete sequence.
ACCESSION	AY569780
VERSION	AY569780.1 GI:45861238
KEYWORDS	
SOURCE	Cloning vector pTCR-Auto
ORGANISM	Cloning vector pTCR-Auto
REFERENCE	other sequences; artificial sequences; vectors.
AUTHORS	1 (bases 1 to 12404)
TITLE	Wu, G., Wang, Y. and Zhu, Z.
JOURNAL	Direct Submission
	Submitted (09-MAR-2004) Fish Genetics, Institute of Hydrobiology,
	Chinese Academy of Science, 7# South Road of East Lake, Wuhan,
	Hubei 430072, People's Republic of China
	Location/Qualifiers
	1. 12404
	/organism="Cloning vector pTCR-Auto"
	/mol_type="other DNA"
	/db_xref="taxon:268734"
	419. 452
	/note="Loxp RE"
	complement (505. 555)
	/note="SV40 polyA"
	complement (707. 1387)
	/codon_start=1
	/transl_table=11
	/product="RFP"
	/protein_id="AAS78500.1"
	/db_xref="GI:45861239"
	/translation="MVRSSKNVKEFMRFYVMEGTVNGHEFEIEGEGGRPYEGHNT
	VKLKVTYKGGPLPFAWDILSPQFQYGSKYVYKHPADIPDYKLSFPEGFKERVMNPF
	GGVVTQDSSLQDGCETIKYKKEFGVNFPSDGPVMQKTMGEASTERLYPRDGLKGG
	EIHKALKLKDGGHYLVEFKSIYMAKKPVQLPGYVVDSKLDITSHNEDYTIIVEQYERT
	EGRHHLFL"
	complement (1447. 2036)
	/note="CMV promoter"
	complement (2387. 2761)
	/codon_start=1
	/transl_table=11
	/product="zeocin"
	/protein_id="AAS78501.1"
	/db_xref="GI:45861240"
	/translation="MAKLTSAPVLTARDVAGAVEFWTDRLGFSRDFVEDDFAGVVRD
	DVTLFISAVQDQVPDNTLAWVWRGLDELYAEWSEVSTNFRDASGPAMTEIGEQPMW
	GREFALRDPAGNCVHFVABEQD"
	complement (2860. 2990)
	/note="T7 terminator"
	complement (3066. 4097)
	/codon_start=1
	/transl_table=11
	/product="Cre"
	/protein_id="AAS78502.1"
	/db_xref="GI:45861241"
	/translation="MANLTVHQNLPAIPVDATSDVRKNLMDMFRDQAFSEHTWKM
	LSVCRSWAAWCKLNNRKMFPAPEDVDVLLYLQARGLAVKTIQCHLQGLMLHRRS
	GLPRPSDSNAVSLVMRIRIKENVDAGERAKOALAFERTDFDQVRSLSMNSDRCODIRN
	LAFLGIAYNTLLRIAEIARIKVDISRTDGRMLIHIGRTKTVSTAGVEKALSLGVT
	KLVERMISVSGVADDPNNYLFCDVRKNGVAAPSATSQLSTRALGCIPEATHRLIYGAK
	DDSGGRYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTNNVIMVYIRNLDSSETGA
	MVRLLEDGD"
	complement (4109. 4683)
	/note="Internal ribosome entry sequence from
	encephalomyocarditis virus"
	complement (4745. 4762)
	/note="T7 promoter"
	complement (4837. 4967)
	/note="T7 terminator"
	complement (5009. 7660)
	/codon_start=1
	/transl_table=11
	/product="T7 RNA polymerase"
	/protein_id="AAS78503.1"
	/db_xref="GI:45861242"
	/translation="MNTINIAKNDFSDIEILAIIPNTLADHYGGRLARQOLALEHESY

```

EMGEARFRKMFEROLKAGEVADNAAAKPLITTLPKMIARINDWFEVVKAKRGKRP
FOFLOEIKPEAVAYITIKTTLACTSADNTTVOAVASAIGRAIEDARFGRIRDL
HEFKNVEEOLNKRVGVHVIYKKAQVVEADMLSKGLLGGEASSNHKEDSIHVGR
MLESTGMVSLHRONAGVVGODSETIELAPEYAEAIATRAGALAGISPMFOCV
PWTGITGGYWANGRRPLALVTRHSSKALMRVEDVYMEVYKAINAONTAKINKV
LAVANVITKMKHCPVEDIPAIERBELPMKPEDIDMNPALTAWKRAAAAVYRKD
SRISLEFMLEOANKFANHKAIWFPYNDMDGRVYAVSMFNPQNDMTKGLITLAK
PCKEGGYWLKIHGANCAGVDKYPPEPERIKFIEENHENIMACAKSPLENTWAEQ
FCFLAFCFEYAGVOHGHLSYNSCLPLAFDGSCEIOHFSAMLEBEGRAVNLPS
VDIYIGIVAKKNEILOADANGTDNEVTVTDENTGEISEKVLGTKALAGWLA
VTRSVTKRSVMTLAYGSKFEPQOVLDTIOPAIDSGKLMFPQNPQAGYMAKL
ESVSVTVVAAVEAMNMLKSAKLAAEVKDKTGEILRKRCARVHVTDPGFVWQ
KPIOTRLNLMFLOFLOPTINMKDSEIDAKHGESIAPNFVSODGSHLRTVYV
HEKYGIESFALIHDSFGTIPADAANLFAVRETMTDVTYESCDVLADFYDQFAD
OLDKMPALPAKGNLNLRLDILSEFAFA"
misc_signal
/note="SV40 T-antigen NLS"
complement(7670..7690)
misc_feature
complement(7702..8276)
/note="internal ribosome entry sequence from
encephalomyocarditis virus"
promoter
/note="T7 promoter"
complement(8338..8355)
CDS
complement(8744..9826)
/codon_start=1
/transl_table=11
/product="LacI"
/protein_id="AAS78504.1"
/db_xref="GI:45861243"
/translation="MKPTLVDAEYAGVSQYQTVSRVNVNQASHVSAKTRKEV
LNYIPNRVAQQLAGKQSLIGVATSSALHAPSQIVAAIKSRADQLGASVVSME
GYEACKAAVHNLIAORVSGLTINPLDODAIAVEACTNVPALFLDSDQPIINSI
FSHEDGTRLGVEHLVALGHQOIALLAGPLSSVSARLRLAGMKYLRTRNOIQPI
DMSAMSGFOQTQMNLNEGIVETAMLVANDOMALGAMRAITESGVRGADSHSV
EDSSCIYIPPTTIKODFRLIGOTSVDRLLQLSQGQAVKGNQLPVSIVKRTT
QTASPRALADSLMQLARQVSRLESQ"
misc_feature
10113..10146
/note="LoxP LE"
10585
/note="Col E1"
complement(11344..12204)
CDS
/codon_start=1
/transl_table=11
/product="Amp"
/protein_id="AAS78505.1"
/db_xref="GI:45861244"
/translation="MSIQHFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGAR
YIELDINGKILLESFRPEERFPMSTFKVLLCGAVLSRIDAGSGLGRIRHYSND
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGSPKELTAFLAHMGD
DRWEPELNEAIPIIDERDTTMVAMATTLRKLTGLLTLASROOLDMEADKVA
LRSALPAGWFIADKSGAGERSRGIITAAIGPDGKPSRIIVITYTGSQATMDER
EIGASLIKHW"
ORIGIN
Alignment Scores:
Pred. No.: 1.24e-112 Length: 12404
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 11 Gaps: 0
US-10-006-922A-12 (1-225) x AY569780 (1-12404)
QY 1 MetArgSerSerLyAsnValIleYsGluPheMetArgPheLybValaArgMetGluGly 20
Db 1384 GTGCGCTCTCCAAAGACGTCAATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 1325
QY 21 ThrValaSnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyGluGly 40
Db 1324 ACCGTGAACGCCACGAGTTCCAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGC 1265
QY 41 HisAsnThrVallybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60

```



Db	1264	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTCGCTGGGACATC	1205
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro	80
Db	1204	CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCAGACATCCCC	1145
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	1144	GACTACAGAAGCTGTCTTCCCTCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	1085
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	1084	GACGGCGCGGTGTGACCGGTGACCACCAAGACTCTCCCTGCAAGACGGCTGCTCATCTAC	1025
QY	121	LysValLysPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	1024	AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCGCTAATGCAGAAAGAAGACC	965
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	964	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGGTGTGAAGGGCGAG	905
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	904	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGAAGTCAAGTCCATC	845
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	844	TACATGGCCCAAGAACCCCGTGACAGTCCCGGCTACTACTACGTGACTCCAAAGCTGGAC	785
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	784	ATCACCTCCCAACAAGAGACTACACCAATCGTGAAGACAGTACGAGCGCACCGAGGGCCGC	725
QY	221	HisHisLeuPheLeu 225	
Db	724	CACCACCTGTTCCTG 710	
RESULT 21			
LOCUS	AX370406	678 bp	DNA linear PAT 16-FEB-2002
DEFINITION	Sequence 3 from Patent WO0196373.		
ACCESSION	AX370406		
VERSION	AX370406.1 GI:18857491		
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS	Fradekov,A.F. and Tersikh,A.		
TITLE	Fluorescent timer proteins and methods for their use		
JOURNAL	Patent: WO 0196373-A 3 20-DEC-2001;		
	Clontech Laboratories Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..678		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="variant of sequence from Discosoma sp."		
ORIGIN			
Alignment Scores:	7.26e-114	Length:	678
Pred. No.:	1207.00	Matches:	223
Score:	99.56%	Conservative:	1
Percent Similarity:	99.11%	Mismatches:	0
Best Local Similarity:	99.11%	Indels:	0
Query Match:	99.42%	Gaps:	0
DB:	6		
US-10-006-922A-12 (1-225) x AX370406 (1-678)			
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20

Db	1	ATGGCTCTCCAGAAGCGTCATCAAGAGTTCAATGCCCTTCAAGGTGGCATGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGAGGGCCCGCTTCAGAGGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGGCCCCCTTGCCCTTCGGTGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCAGACATCCCC	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGCGGTGGCGGACCGGTGACCCAGACTCTCCCTGCAGAGACGGCTGCTCATCTAC	360
QY	121	LysValLysPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCGCTGATGCAGAGAAGACC	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGGTGTGAAGGGCGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGAAGTCAAGTCCATC	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCCAAGAACCCCGTGACAGTCCCGGCTACTACTACGTGAAGACCAAGCTGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACCTCCCAACAAGAGACTACACCATCGTGAAGACAGTACGAGCGCACCGAGGGCCGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCACCTGTTCCTG 675	
RESULT 22			
LOCUS	AX348043	666 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 22 from Patent EP1172375.		
ACCESSION	AX348043		
VERSION	AX348043.1 GI:18614153		
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS	Odenthal,M. and Jung,D.		
TITLE	Gene expression, genome alteration and reporter gene expression in myofibroblasts and myofibroblasts-like cells by using the regulatory regions within the alpha smooth muscle alpha-actin gene		
JOURNAL	Patent: EP 1172375-A 22 16-JAN-2002;		
	Odenthal, Margarete (DE)		
FEATURES	Location/Qualifiers		
source	1..666		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="Red Fluorescent Protein"		
ORIGIN			
Alignment Scores:	9.44e-113	Length:	666
Pred. No.:			
QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20



Score:	1196.00	Matches:	221
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.52%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) X AX348043 (1-666)

QY	5	LYSASNVALILELYSGIUPHEMELARGPHELYSVALARGMETGLYLTHRVALLAENGLY	24
DB	1	AAGAATGTTATCAAGGAGTTCAATGAGGTTTAAAGGTTCCGATGGAAAGAACCGTCAATGGG	60
QY	25	HISGLUPHEGLUILEGGLYGLUGLYGLUGLYARGPROTYRGLYGHISASNTHRVAL	44
DB	61	CACGAGTTTGAATAGAAAGCCGAGAGAGAGGAGGCCATACGAAGGCCAATAACCGTA	120
QY	45	LYSLEULYSVALTHRLYSGLYGLYPROLEUPROHEALATRPASPILEUSERPROGLN	64
DB	121	AAGCTTAAGGTAACCAAGGGGGACCTTGGCCATTGGCTTGGATATTTGTCAACCAAA	180
QY	65	PHENGLNTRYGLYSERLYSVALTYRVALLYSHISPROALAASPILEPROASPTYRLYSLYS	84
DB	181	TTTCAGTATGGAAGCAAGGTATATGCAAGCAACCTGCCGACATACCAACTATAAAAG	240
QY	85	LEUSERPHEPROGLUGLYPHELYSTRPGLUARGVALMETASNPHENGLUASPGLYGLYVAL	104
DB	241	CTGTCAATTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAAGACGGTGGCGTC	300
QY	105	VALTHRVALTHRGLNASPSEERSEULEGLNASPGLYCYSPHEILETYRLYSVALLYSPHE	124
DB	301	GTTACTGTAAACCCAGGATTCAGTTTGCAAGATGGCTGTTTCATCTACAGGTCAAGTTTC	360
QY	125	IIEGLYVALASNPHEPROSEERASPGLYPROVALMETGLNLYSLYSTHMETGLYTRPGLU	144
DB	361	ATTGGCGTGAACCTTCCTCCGATGGAACCTGTTATGCAAAAGAACAATGGCGTGGGA	420
QY	145	ALASERTHRGLVARGLEUTYRPROARGASPGLYVALLEULYSGLYGLUILEHISLYSALA	164
DB	421	GCCAGCACTGAGCGTTTGATCTCGTGATGGCGTGTTGAAGAAGAGATTCATAAGGCT	480
QY	165	LEULYSLEULYSASPGLYGLYHISTYRLEUVALGLUPHELYSERTIETYRMETALALYS	184
DB	481	CTGAAGCTGAAAAGACGGGTGTCATTACTAGTTGAATTCAAAAGTATTTACATGGCAAAG	540
QY	185	LYSPROVALGLNLEUPROGLYTYRTYRVALASPSERLYSLEUASPILETHRSERHIS	204
DB	541	AAGCCTGTGACGCTACCAAGGTACTACTATGTTGACTCCAAACTGGAATATAACAAGCCAC	600
QY	205	ASNGLUASPTYRTHRIIEVALIGLUGINTYRGLUARGTHRGLUGLYARGHISHSLEUPHE	224
DB	601	AACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGGCCAACCATCTGTTC	660
QY	225	LEU 225	
DB	661	CTT 663	

RESULT 23					
AX353910	AX353910	666 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX353910				
DEFINITION	Sequence 22 from Patent WO0204509.				
ACCESSION	AX353910				
VERSION	AX353910.1	GI:18618888			
KEYWORDS	.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				
	other sequences; artificial sequences.				

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
1	Odenthal, M. and Jung, D.	Gene expression, genome alteration and reporter expression in myofibroblasts and myofibroblast-like cells	Patent: WO 0204509-A 22 17-JAN-2002;	
	Odenthal, Margarete (DE)		Location/Qualifiers	

```

1. .666
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Red Fluorescent Protein"

```

### Alignment Scores:

Pred. No.:	9.44e-113	length:	666
Score:	1196.00	Matches:	221
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.52%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AX353910 (1-666)

QY		5 LysAsnValIleLysGluPheMetArgPheLysValArgMetCylGlyThrValAsnGly	24
Dd		1 AAGAATGTATTCAAGSAGTTTCATGAGGTTTAAGGTTCCGCATGGAAAGCGTCAATTGGG	60
QY		25 HisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrgLuglYHisAsnThrVal	44
Dd		61 CACGAGTTTGAAATGGAAGGCCGAAGAGAGAGGGAGGCCCATACGAAGGCCCAATAACCGTA	120
QY		45 LysLeuLysValThrLysGlyGlyPProLeuProPheAlatTPAspIleLeuSerProGln	64
Dd		121 AAGCTTAAGGTAAACCAGAGGGGGACCTTGCCATTGGCTTGGAATATTTGTGCACCACAA	180
QY		65 PheGlnTyrgLysSerLysValTyrrValLysHisProAlaAspIleProAspTyrlLysLys	84
Dd		181 TTTCAGTAGGAGCAAGGATATATGTCAAGCACCTCGCCGACATACCAACTATAAAAAAG	240
QY		85 LeuSerPheProGluGlyPheLysTrpGluaArgValMetAsnPheGluAspGlyGlyVal	104
Dd		241 CTGTCAATTTCCCTGAAGCATTTAATGGGAAGGCTCATGAACITTTGAAGACGGTGGCGTC	300
QY		105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrlLysValLysPhe	124
Dd		301 GTTACTGTAAACCAGGATTCACGTTTGCAAGATGGCTGTTTCATCTACAAGGTCAAAGTTTC	360
QY		125 IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu	144
Dd		361 ATTGGCGTGAACCTTCCCTCCGATGGAACCTGTATGCAAAAGAAGCAATGGCGTGGGA	420
QY		145 AlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLysGlyLuiLeHisLysAla	164
Dd		421 GCCAGCAGTGAAGCGTTTGTATCCTCGTATGGCGCTGTGAAAAGAGAGATTTCATAAGGCT	480
QY		165 LeuLysLeuLysAspGlyGlyHisTyrrLeuValGluPheLysSerIleTyrrMetalalys	184
Dd		481 CTGAAGCTGAAGAAGCGGTGTCTATTACCTAGTTGAATTCAAAAGTATTTCATATGGCAAAAG	540
QY		185 LysProValGlnLeuProGlyTyrrTyrrValAspSerLysLeuAspIleThrSerHis	204
Dd		541 AAGCCTGTGACGTAACCAAGGTACTACTATGTGACTCCAACCTGGATATAACAAGCCAC	600
QY		205 AsnGluAspTyrrThriLeValGluGlnTyrgLuArgThrGluGlyArgHisHisLeuPhe	224
Dd		601 AACGAAGACTATACATCGTTGAGCAGTATGAAGAAGAACCGAGGCCACCATCTGTTC	660
QY		225 Leu 225	
Dd		661 CTT 663	

RESULT 24	
LOCUS	AY679107
DEFINITION	AY679107 711 bp mRNA linear INV 25-AUG-2004 Discosoma sp. RC-2004 enhanced red fluorescent protein R+ mRNA, complete cds.
ACCESSION	AY679107
VERSION	AY679107.1 GI:51472046
KEYWORDS	.



SOURCE Discosoma sp. RC-2004  
ORGANISM Discosoma sp. RC-2004  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE 1 (bases 1 to 711)  
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.  
TITLE Cloning of Anthozoan Fluorescent Protein Genes  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 711)  
AUTHORS Carter, R.W., Gibbs, P.D.L. and Schmale, M.C.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker  
Cswy, Miami, FL 33129, USA  
FEATURES  
source  
1..711  
location/Qualifiers  
/organism="Discosoma sp. RC-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:289055"  
1..711  
/note="GFP-like; DsRed+"  
/codon\_start=1  
/product="enhanced red fluorescent protein R+"  
/protein\_id="AAU04444.1"  
/db\_xref="GI:51472047"  
/translation="MSSCKNVIKFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTV  
KLKVTKGGLPFAWDILSPQYGSKVYVKHPADIPYKKLSPEGFKWERVMNPEDG  
GVTVTQDSSLQDGCFTYKVKFIGNVPSDGPVMQKTMGWEASTERLYPRDGYLKGE  
IHKALKLDGGHYLVFEFKTIYMAKKPVQLPGYYVDSKLDITSHNKDYTIIVEQYERTE  
GRHHLFLKABLGSNNGER"  
ORIGIN

Alignment Scores:  
Pred. No.: 1.02e-112 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-12 (1-225) x AY679107 (1-711)

QY 1 MetArgSerSerLySAbnValIleLySGluPheMetArgPheLyValAlaArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAGAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGTATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGAGAGAGAGGAGCCATACGAAGGC 120  
QY 41 HisAsnThrValIleLeuLySValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGAGCCTTGGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValIleHisIleProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCATCCTGCCGACATACCA 240  
QY 81 AspTyrIleValLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAGCTGTCATTCTCGAAGGATTTAATGGGAAAGGGTCATGAACCTTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTGTAAGTAAACCAAGATTCAGTTTGACAGATGGCTGTTTCATCTAC 360  
QY 121 LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleLySThr 140  
DB 361 AAGGTCAAGTTCATGGCGTGAACCTTCTCTGATGACCTGTATATGCAAAAGAAAGACA 420  
QY 141 MetGlyTrpGlyAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCTCTGATGGCGTGTGAAGAAGAGAG 480

QY 161 IleHisLySAleuLyLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 481 ATTCAATAAGGCTCTGAAGTTGAAAGACGGTGTGCTTACCTAGTTGAATTCAAACTAATT 540  
QY 181 TyrMetAlaLySProValGlnLeuProGlyTyrTyrTyrValAlaAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAAGCCTGTGCAAGCTACCAAGGGTACTACTATGTGACTCCAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATACAAGCCACACAACAAGACTATACAATCGTTGACAGATATGAAGAACCAGAGGAGCC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTCTCTT 675

RESULT 25  
AF545828 678 bp mRNA linear INV 04-JUN-2004  
LOCUS AF545828  
DEFINITION Discosoma sp. JW-2002 orange fluorescent protein FP586 mRNA,  
complete cds.  
ACCESSION AF545828  
VERSION AF545828.1 GI:33333763  
KEYWORDS  
SOURCE Discosoma sp. JW-2002  
ORGANISM Discosoma sp. JW-2002  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE 1 (bases 1 to 678)  
AUTHORS Wiedemann, J. and Girod, A.  
TITLE Direct Submission  
JOURNAL Submitted (12-SEP-2002) Abteilung Allgemeine Zoologie und  
Endokrinologie, Universitaet Ulm, Albert Einstein Allee 11, Ulm  
89069, Germany

FEATURES  
source  
location/Qualifiers  
1..678  
/organism="Discosoma sp. JW-2002"  
/mol\_type="mRNA"  
/db\_xref="taxon:208461"  
/note="Isolated from specimen with smooth disc"

CDS

1..678  
/note="GFP-like protein, orange-red; dsFP586; similar to  
fluorescent protein drFP583 (dsRed)"  
/codon\_start=1  
/product="orange fluorescent protein FP586"  
/protein\_id="AAQ11987.1"  
/db\_xref="GI:33333764"  
/translation="MSSCKNVIKFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTV  
KLKVTKGGLPFAWDILSPQYGSKVYVKHPADIPYKKLSPEGFKWERVMNPEDG  
GVTVTQDSSLQDGCFTYKVKFIGNVPSDGPVMQKTMGWEASTERLYPRDGYLKGD  
IHKALKLDGGHYLVFEFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERTE  
GRHHLFL"

ORIGIN

Alignment Scores:  
Pred. No.: 1.95e-112 Length: 678  
Score: 1193.00 Matches: 221  
Percent Similarity: 98.67% Conservative: 1  
Best Local Similarity: 98.22% Mismatches: 3  
Query Match: 98.27% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922A-12 (1-225) x AF545828 (1-678)

QY 1 MetArgSerSerLySAbnValIleLySGluPheMetArgPheLyValAlaArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAGAATGTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGTATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGAGAGAGGAGCCATACGAAGGC 120  
QY 41 HisAsnThrValIleLeuLySValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60



Db	121	CAACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGTGGGATATT	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	TTGTCAACCACAATTTTCAATATGGAAGCAAGGATATATGTCAGACACCCTGCCACATACCA	240
Qy	81	AspTyrLysLysLeuSerPheProGlyGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	241	GACTATATAAAAGCTGTTCATTTCTGAAAGATTAAATGGGAAAGGTCATGAACCTTGAA	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGTGGCGTCTTACTGTAAACCAAGATTCCAGTTTGAGAGATGGCTGTTTCATCTAC	360
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGTCAAGTTTCATTGGCGGTGAACCTTCTCTGATGGACCTGTTATGCAAAAGAAGACA	420
Qy	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyIu	160
Db	421	ATGGGCTGGGAAGCCAGCAGCTGAGCGTTTGTATCCTCGTGATGGCGGTGTGAAGAGAT	480
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTCAATAAGGCTCTGAAGCTGAAGAAGCGTGTTCATTACCTAGTTGAATTCAAAAGTATT	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCAAGAAGACCTGTGTGACGTAACCAAGGGTACTATGTTGACTCCAAACTGGAT	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAAGAACGAGGAGCG	660
Qy	221	HisHisLeuPheLeu 225	
Db	661	CACCATCTGTTCCTT 675	
RESULT 26			
LOCUS	AX370408	678 bp	DNA linear PAT 16-FEB-2002
DEFINITION	Sequence 5 from Patent WO0196373.		
ACCESSION	AX370408		
VERSION	AX370408.1 GI:18857492		
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 Pradkov, A.F. and Tersikh, A.		
TITLE	Fluorescent timer proteins and methods for their use		
JOURNAL	Patent: WO 0196373-A 5 20-DEC-2001;		
	Clontech Laboratories Inc. (US)		
FEATURES			
source			
	1.678		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="variant of sequence from Discosoma sp."		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.13e-112	Length:	678
Score:	1191.00	Matches:	220
Percent Similarity:	98.67%	Conservative:	2
Best Local Similarity:	97.78%	Mismatches:	3
Query Match:	98.11%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AX370408 (1-678)			
Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyIu	20

[illegible]



Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX824732 (1-678)

```
QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
        |||:::|||||
DB      1 ATGGCTCTCTCCGAGAACGTGATCAACGAGTTGCGCTCAAGGTGCGCATGAGGCG 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
        |||:::|||||
DB      61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC 120

QY      41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60
        |||:::|||||
DB      121 CACAACACCGTGAAGTTGAAGGTGACCAAGGGCGGCCCTCGCTTCCGCTGGGACATC 180

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
        |||:::|||||
DB      181 CTGTCCCCCAAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCAGCATCCCC 240

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
        |||:::|||||
DB      241 GACTACAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 300

QY      101 ArgGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        |||:::|||||
DB      301 GACGGCGGCGTGGCGACCGGTGACCCAGGACTCTCTCCGTGAGGACGGCTGCTCATCTAC 360

QY      121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
        |||:::|||||
DB      361 AAGGTGAAGTTGATCGCGGTGAACCTTCCCTCCGACGGCCCCCGTGATGCAGAAAGAGACC 420

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
        |||:::|||||
DB      421 ATGGGCTGGGAGGCTCCACCGGACGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 480

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
        |||:::|||||
DB      481 ATCCACAAGGCCCTGAAGCTGAAGAGACGGGGCCACTACCTGGTGAAGTTCAAGTCCATC 540

QY      181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
        |||:::|||||
DB      541 TACATGGCCCAAGAGCCCGGTGACGCTGCGGCTACTACTACGTGACCAAGCTGAGAC 600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
        |||:::|||||
DB      601 ATCACCTCCCAACAAGGAGACTACACCATGTGAGAGCACTACGAGCGCACCGAGGGCCGC 660

QY      221 HisHisLeuPheLeu 225
        |||:::|||||
DB      661 CACCACTGTCTCTG 675

RESULT 28
AY679106          921 bp      mRNA      linear      INV 25-AUG-2004
AY679106          Discosoma sp. RC-2004 red fluorescent protein R1 mRNA, complete
LOCUS             AY679106
DEFINITION        cds.
ACCESSION         AY679106
VERSION           AY679106.1
KEYWORDS          Discosoma sp. RC-2004
SOURCE            Discosoma sp. RC-2004
ORGANISM          Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
                  Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE         1 (bases 1 to 921)
AUTHORS           Carter,R.W., Gibbs,P.D.L. and Schmale,M.C.
TITLE             Cloning of Anthozoan Fluorescent Protein Genes
JOURNAL           Unpublished
REFERENCE         2 (bases 1 to 921)
AUTHORS           Carter,R.W., Gibbs,P.D.L. and Schmale,M.C.
TITLE             Direct Submission
```

JOURNAL Submitted (06-JUL-2004) MBF, University of Miami, 4600 Rickenbacker  
Cswy, Miami, FL 33129, USA  
FEATURES Location/Qualifiers  
source 1..921

1..921  
/organism="Discosoma sp. RC-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:289055"  
1..66  
67..777  
/note="GFP-like; DsPr1"  
/codon\_start=1  
/product="red fluorescent protein R1"  
/protein\_id="AAU0443.1"  
/db\_xref="GI:51472045"

/translation="MSSCSKNVIEKFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTV  
KLKVTKGGLPPAMDILSPQFGYSKVYVGHPIIDPYKGLSFEGRFKERVMNFEDG  
GVTVTQDPSLDGCFIYKVKFIVGNFSPDGPVMOKTMGWEASTERLYPRDGLKGE  
IHKALKDKDGHYLVFEFKTIYMAKKPVQLPGYYVDSKLDITSHNKDYTLVEQYERTE  
GRHHLFLKAEISNVGER"

3'UTR  
polya\_signal  
ORIGIN

Alignment Scores:  
Pred. No.: 4.54e-112 Length: 921  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0  
DB: 2 Gaps: 0

US-10-006-922a-12 (1-225) x AY679106 (1-921)

```
QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
        |||:::|||||
DB      67 ATGAGTTGTTCCAGAAGATGTATCAAGAGTTTCATGAGGTTTAAAGTTCGTATGGAAGGA 126

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
        |||:::|||||
DB      127 ACGGTCAATGGGCACGAGTTGAAATAGAAGCGCAAGAGAGAGGACCATACGAAGGC 186

QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60
        |||:::|||||
DB      187 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTGGCATTTGCTTGAGATATT 246

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
        |||:::|||||
DB      247 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCATCTCGCGACATACCA 306

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
        |||:::|||||
DB      307 GACTATAAAAGCTGTGATTTCTTGAAGATTAAATGGGAAGGCTCATGAACCTTGAA 366

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        |||:::|||||
DB      367 GACGTTGGCTCGTTACTGTAAACCAAGATCCAGTTGCAGAGATGTTGTTCAATTAC 426

QY      121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
        |||:::|||||
DB      427 AAGGTCAAGTTGATGGCGTGAACCTTCTTGATGAGACCTGTTATGCAAAAAGAGACA 486

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
        |||:::|||||
DB      487 ATGGGCTGGGAAGCCACGACTGAGCGTTTGTATCTCTGATGAGCGCTGTGAAGAAGAGAG 546

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
        |||:::|||||
DB      547 ATTCAATAGGCTCTGAAGTTGAAGAAGCGGTGTATTAACCTTAAGTTCAAAACTATT 606

QY      181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
        |||:::|||||
DB      607 TACATGGCAAGAAGCCTGTGACGACTACAGGGTACTACTATGTTGACTCCAAAAGTGAT 666
```



Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	667	ATAACAGACCACAACAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGGAGCG	726
Qy	221	HisIleuPheLeu	225
Db	727	CATCATCTGTTCTT	741
RESULT 29			
AX824731		675 bp	DNA
LOCUS	AX824731		linear
DEFINITION	Sequence 13 from Patent WO02068459.		PAT 11-DEC-2003
ACCESSION	AX824731		
VERSION	AX824731.1	GI:39750594	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1	other sequences; artificial sequences.	
AUTHORS			
TITLE	Non aggregating fluorescent proteins and methods for using the same		
JOURNAL	Patent: WO 02068459-A 13 06-SEP-2002;		
FEATURES	location/Qualifiers		
SOURCE	1..675		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="non-aggregating mutant"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.01e-111	Length:	675
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AX824731 (1-675)			
Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGGCCTCCTCCGAGAACGTATCAACGAGTTCAATGCGCTTCAAGTGGCATGAGGGC	60
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC	120
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	121	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTTGCCTGGACATC	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAACACCCCGGCACATCCC	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	241	GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGCGCTGGCGACCGTGACCCAGSACTCTCTCCCTGCAAGAGCGGCTTCACTAC	360
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAAGTGAAGTTCAATCGCGCTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAAGAAC	420
Qy	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGAGCGGCTGTGAAGGGCGAG	480
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180

Db	481	ACCCACAAGGCCCTGAGCTGAAGGACGCGGCCCACTACTGCTGTGAGTTCAAGTTCATC	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCCCAAGCTGAC	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACCTCCACAACAGGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGCCGC	660
Qy	221	HisIleuPheLeu	225
Db	661	CACCACCTGTCTCTG	675
RESULT 30			
AX666133		1050 bp	DNA
LOCUS	AX666133		linear
DEFINITION	Sequence 7 from Patent WO02060941.		PAT 26-MAR-2003
ACCESSION	AX666133		
VERSION	AX666133.1	GI:29290961	
KEYWORDS			
SOURCE	unidentified		
ORGANISM	unidentified		
REFERENCE	1	unclassified sequences.	
AUTHORS			
TITLE	Zhao, M., Xu, M., Jiang, P. and Yang, M.		
JOURNAL	Fluorescent proteins		
	Patent: WO 02060941-A 7 08-AUG-2002;		
	Anticancer, Inc. (US)		
FEATURES	location/Qualifiers		
source	1..1050		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	/note="Coral"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.73e-111	Length:	1050
Score:	1186.00	Matches:	219
Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x AX666133 (1-1050)			
Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	289	ATGGCCTCCTCCGAGAACGTATCAACGAGTTCAATGCGCTTCAAGTGGCATGAGGGC	348
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	349	ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC	408
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	409	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTTGCCTGGACATC	468
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	469	CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCC	528
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	529	GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	588
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	589	GACGGCGCGTGGCGACCGTGACCCAGAGACTCTCTCCCTGCAAGAGCGGCTTCACTAC	648
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140



Db 649 AAGGTGAAGTTCATCGCGCTGAAGTTCCTCCGACGGCCCCCGATGCAGAGAAGACC 708  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 709 ATGGGCTGGAGGCGCTCCACCGAGCGCCTGTACCCTCCGACGGCGTGCTGAAGGGCGAG 768  
Qy 161 HLeHsAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 769 ACCGACAAGGCCCTGAAGCTGAAGGACGGCGGCACCTACCTGTGAGGTCAAGTCCATC 828  
Qy 181 TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 829 TACATGGCCAGAAGACCCGCTGACAGCTGCCGCTACTACTACGTGAGCGCCAAAGCTGAC 888  
Qy 201 HLeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLysArg 220  
Db 889 ATCACCCTCCCAACAGAGGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCCGC 948  
Qy 221 HsHsIleuPheLeu 225  
Db 949 CACCACCTGTTCTCTG 963

RESULT 31  
CQ849509 4488 bp DNA linear PAT 23-AUG-2004  
LOCUS Sequence 6 from Patent WO2004067751.  
DEFINITION CQ849509  
ACCESSION CQ849509  
VERSION CQ849509.1 GI:51507513

KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
AUTHORS  
TITLE  
1 Coullard-Despres,S., Karl,C., Kuhn,H.G. and Aigner,L.  
Use of regulatory sequences for specific, transient expression  
in neuronal determined cells

JOURNAL  
Patent: WO 2004067751-A 6 12-AUG-2004;  
Klinikum der Universitaet Regensburg (DE)

FEATURES  
source  
1..4488  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="fusion construct homo sapiens and DsRed2"

ORIGIN

Alignment Scores:  
Pred. No.: 1.01e-110 Length: 4488  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CQ849509 (1-4488)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 3572 ATGGCCTCCTCCGAGAACGTTCATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG 3631  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 3632 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGCG 3691  
Qy 41 HisAsnThrValLysLeuValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 3692 CACCAACACCGTGAAGCTGAAGTGAACCAAGGGCGGCGCCCTTCGCTGAGCATC 3751  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 3752 CTGTCCCCCAGTTCAGTACGCGCTCCAAGGTGTACGTGAAGCACCCCGCGCATCCCC 3811

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTrpGluArgValMetAsnPheGlu 100  
Db 3812 GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 3871  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 3872 GACGGCGCGCTGGCGGACCGGTGACCCAGGACTCTCCCTGCAGAGACGGCTGCTTCATCTAC 3931  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 3932 AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGGCGCCCTGTATGCAGAGAAGACC 3991  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 3992 ATGGGCTGGAGGCGCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 4051  
Qy 161 HLeHsAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 4052 ACCGACAAGGCCCTGAAGCTGAAGGACGGCGGCACCTACCTGTGAGTTCAAAGTCCATC 4111  
Qy 181 TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 4112 TACATGGCCAGAAGACCCGCTGACAGCTGCCGCTCCGCTACTACTACGTGAGCGCCAAAGCTGAC 4171  
Qy 201 HLeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLysArg 220  
Db 4172 ATCACCCTCCCAACAGAGGACTACCATCGTGAGCAGTACGAGCGCACCGAGGGCCGC 4231  
Qy 221 HsHsIleuPheLeu 225  
Db 4232 CACCACCTGTTCTCTG 4246

RESULT 32  
AY818375 4546 bp DNA circular SYN 12-APR-2005  
LOCUS Cloning vector pSAT6-DsRed2-C1, complete sequence.  
DEFINITION AY818375  
ACCESSION AY818375  
VERSION AY818375.1 GI:56553574

KEYWORDS  
SOURCE  
ORGANISM  
Cloning vector pSAT6-DsRed2-C1  
Cloning vector pSAT6-DsRed2-C1  
other sequences; artificial sequences; vectors.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 4546)  
Tzfira,T., Tian,G.W., Lacroix,B., Vyas,S., Li,J., Leitner-Dagan,Y.,  
Krichevsky,A., Taylor,T., Vainstein,A. and Citovsky,V.  
pSAT vectors: a modular series of plasmids for autofluorescent  
protein tagging and expression of multiple genes in plants

JOURNAL  
Plant Mol. Biol. 57 (4), 503-516 (2005)

FEATURES  
PUBMED 15821977  
2 (bases 1 to 4546)  
Tzfira,T., Lacroix,B. and Citovsky,V.  
Direct Submission  
Submitted (05-NOV-2004) Department of Biochemistry and Cell  
Biology, State University of New York Stony Brook, 100 Nicolls Rd,  
Stony Brook, NY 11794, USA

location/Qualifiers  
1..4546  
/organism="Cloning vector pSAT6-DsRed2-C1"  
/mol\_type="other DNA"  
/db\_xref="taxon:301574"  
441..762  
/note="CMV 35S promoter"  
768..1089  
/note="CMV 35S promoter"  
1190..1320  
/note="5' UTR from tobacco etch virus; translational  
enhancer"  
1323..1397  
/note="N-terminal end of DsRed2 protein; part of fusion  
protein"  
1398..2066  
/note="multiple cloning site; MCS"  
2067..2277  
terminator



CDS		/note="CamV 35S terminator" complement(3486..4346) /note="amp resistance" /codon_start=1 /product="bla" /protein_id="AAV97912.1" /db_xref="GI:5653575" translation="MSIQHFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGARVGY IELDLNSGKILSEFRPEERFPMWSTFKVLLCGAVLSRIDAGQEQLGRIRHYSQNDLVE YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGSPKELTAFIHNMGDHTRL DRWEPELNEAIPNDERDTTMPVAMATTIRKLLTGSELLTLASRQQLIDWMEADKVA GPLRSALPAGWFIADKSGAGERSRGIITAIAGPDGKPSRIVVIYTTGSAQTMDERNRQIA EIGASLIKHW"			
ORIGIN					
Alignment Scores:					
Pred. No.:	1.03e-110	Length:	4546		
Score:	1186.00	Matches:	219		
Percent Similarity:	98.22%	Conservative:	2		
Best Local Similarity:	97.33%	Mismatches:	4		
Query Match:	97.69%	Indels:	0		
DB:	11	Gaps:	0		
US-10-006-922A-12 (1-225) x AY818375 (1-4546)					
QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly	20		
DB	1323	ATGGCCTCCTCCGAGAACGTCAACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG	1382		
QY	21	ThrValAsnGLYHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40		
DB	1383	ACCGTGAACGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC	1442		
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60		
DB	1443	CACACACCGTGGAAGCTGAAGGTACCAAGGGCGGCCCTCGCTTGGGACATC	1502		
QY	61	LeuSerProGluPheGluTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80		
DB	1503	CTGTCCCCCAGTTCCAGTACGGCTTCCAAGGTGTACGTGAAGCACCCCGCACATCCCC	1562		
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu	100		
DB	1563	GACTACAAAGACTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	1622		
QY	101	AspGlyGlyValValThrValThrGluAspSerSerLeuGlnAspGlyCysPheIleTyr	120		
DB	1623	GACGGCGGCGTGGCGACCGTGACCAAGACTCCTCCCTGCAAGACGGCTGCTCATCTAC	1682		
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140		
DB	1683	AAGTGAAGTTCATCGGCGGTGAACCTCCCTCCGACGGCCCCGTGATGAGAGAAGACC	1742		
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160		
DB	1743	ATGGGCTGGGAGGCGCTCCACCAGCGCCTGTACCCTCCGCGACGGCGTGTGAAGGGCGAG	1802		
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySSerIle	180		
DB	1803	ACCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGGAGTTCAAGTTCATC	1862		
QY	181	TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200		
DB	1863	TACATGGCCAAAGAGCCCGTGCACTGCCCCGGCTACTACTACGTGGAAGCCAAAGCTGAGC	1922		
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220		
DB	1923	ATGACCTCCCAACAAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCGC	1982		
QY	221	HisHisLeuPheLeu	225		
DB	1983	CACCACCTGTTCTCTG	1997		

RESULT 33		DQ005468		4555 bp		DNA	circular SYN 07-MAY-2005
LOCUS		DQ005468					
DEFINITION		Cloning vector pSAT6A-DsRed2-N1, complete sequence.					
ACCESSION		DQ005468					
VERSION		DQ005468.1		GI:63002530			
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
promoter							
promoter							
enhancer							
misc_feature							
CDS							
terminator							
CDS							
ORIGIN							
Alignment Scores:							
Pred. No.:							
Score:							
Percent Similarity:							
Best Local Similarity:							
Query Match:							
DB:							
US-10-006-922A-12 (1-225) x DQ005468 (1-4555)							
QY							
1380							
ATGGCCTCCTCCGAGAACGTCAACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG							
1439							



Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGlyGluGlyArgProTyrGluGly 40  
Db 1440 ACCGTGAACGGCCACGAGTTCGAGTCGAGGGCGAGGGCGGCCCTACGAGGGC 1499  
Qy 41 HisAsnThrValIleuLeuValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 1500 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCTCGCTGGACATC 1559  
Qy 61 LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80  
Db 1560 CTGTCCTCCCGAGTTCAGTACGGCTCCAGGTGTACGTGAAGCACCTCCCGACATCCCC 1619  
Qy 81 AspTyrIleValIleuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu 100  
Db 1620 GACTACAGAAGCTGTCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAG 1679  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 1680 GACGGCGCGTGGCGACCGTGAACCGAGACTCTCCCTGCAGAGCGCTGCTTCATCTAC 1739  
Qy 121 LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleLysThr 140  
Db 1740 AAGGTGAAGTTCAATCGCGGTGAATTCTCCCTCCGACGGCCCGTGATGCAAGAAGACC 1799  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1800 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTCTGAAGGGCGAG 1859  
Qy 161 IleHisIleValIleuLysLeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIle 180  
Db 1860 ACCCACAGGCGCTGAAGCTGAAGAGCGGGCGCACTACCTGTGTGAGTTCAAGTCCATC 1919  
Qy 181 TyrMetAlaIleLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1920 TACATGGCCAGAGACCCCGTGCAGCTGCCGGCTACTACTACGTGAGCGCAAGCTGAGC 1979  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1980 ATCACTCTCCCAACAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 2039  
Qy 221 HisHisLeuPheLeu 225  
Db 2040 CACCACCTGTCTCTG 2054

RESULT 34  
AY818373 4570 bp DNA circular SYN 12-APR-2005  
LOCUS AY818373  
DEFINITION Cloning vector pSAT6-DsRed2-N1, complete sequence.  
ACCESSION AY818373  
VERSION AY818373.1 GI:56553569  
KEYWORDS Cloning vector pSAT6-DsRed2-N1  
SOURCE Cloning vector pSAT6-DsRed2-N1  
ORGANISM Cloning vector pSAT6-DsRed2-N1  
REFERENCE 1 (bases 1 to 4570)  
AUTHORS Tzfira,T., Tian,G.W., Lacroix,B., Vyas,S., Li,J., Leitner-Dagan,Y.,  
Krichevsky,A., Taylor,T., Vainstein,A. and Citovsky,V.  
TITLE pSAT vectors: a modular series of plasmids for autofluorescent  
protein tagging and expression of multiple genes in plants  
JOURNAL Plant Mol. Biol. 57 (4), 503-516 (2005)  
PUBMED 15821977  
REFERENCE 2 (bases 1 to 4570)  
AUTHORS Tzfira,T., Lacroix,B. and Citovsky,V.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2004) Department of Biochemistry and Cell  
Biology, State University of New York Stony Brook, 100 Nicolls Rd,  
Stony Brook, NY 11794, USA  
FEATURES  
source location/Qualifiers  
1. 4570  
/organism="Cloning vector pSAT6-DsRed2-N1"  
/mol\_type="other DNA"  
/db\_xref="taxon:301572"

promoter 441..762  
/note="CaMV 35S promoter"  
promoter 768..1090  
/note="CaMV 35S promoter"  
misc\_feature 1190..1320  
/note="5' UTR from tobacco etch virus; translational  
enhancer"  
misc\_feature 1323..1394  
/note="multiple cloning site; MCS"  
CDS 1395..2072  
/codon\_start=1  
/product="DsRed2"  
/protein\_id="AAV97910.1"  
/db\_xref="GI:56553571"  
/translation="MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTV  
KLKVTKGSPLPFANDILSPQFGYSKVYVKHPADIPDYKLSFPEGFKWERVMNPDG  
GVATVTQDSSLQDGCFTYKVKFIVGNFPSPDGVNQKTMGEASTERLYPRDGLKGE  
THKALKDKDGGHYLVFPKSIYMAKKPVQLPGYTVDAKLDITSHNEDYTIVEQYERTE  
GRHHLFL"  
terminator 2091..2301  
/note="CaMV 35S terminator"  
CDS complement(3510..4370)  
/note="amp resistance"  
/codon\_start=1  
/product="bla"  
/protein\_id="AAV97909.1"  
/db\_xref="GI:56553570"  
/translation="MSIQHFRVALIIPFAAFCLPYFAHPETLVKVKDAEDQLGARVGY  
IELDNSGKILESFRPERFPMSTFKVLLCGAVLSRIDAGEQLGRRIHYSQNDLVE  
YSPVTEKHLLTDGMTVRELCSAAITMSDNTAANLLTTIGSPKELTAFIHNMGDHYTRL  
DRWEPELNEAIPNDERDTMPVAMATTLRKLLTGLLTLASRQQLIDMEADKYAGPL  
LRNALPAGWFIADKSGAGRGSRGIILALGPDGKSRIRIVITYTGSQATMDERNRQIAL  
EIGASLIKHW"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.04e-110 Length: 4570  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AY818373 (1-4570)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheIleValArgMetGluGly 20  
Db 1395 ATGGCCTCTCCGAGAACGTCAATCACCGAGTTCATGCGCTTCAAGGTGGCATGAGGGC 1454  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGlyGluGlyArgProTyrGluGly 40  
Db 1455 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 1514  
Qy 41 HisAsnThrValIleuLysLeuValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 1515 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCCTCGCTGGACATC 1574  
Qy 61 LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro 80  
Db 1575 CTGTCCTCCCGAGTTCAGTACGGCTCCAGGTGTACGTGAAGCACCGCGGACATCCCC 1634  
Qy 81 AspTyrIleValIleuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu 100  
Db 1635 GACTACAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTCGAG 1694  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 1695 GACGGCGCGTGGCGACCGTGAACCGAGACTCTCCCTGCAGAGCGGCTGCTTCATCTAC 1754  
Qy 121 LysValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleLysThr 140  
Db 1755 AAGGTGAAGTTCAATCGCGGTGAATTCCTCCCGACGGCGCGTGTATGCAAGAAGACC 1814



Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
AY640628/c  
LOCUS AY640628 4829 bp DNA circular SYN 01-JUL-2005  
DEFINITION siRNA vector pSUPER-hSyn-DsRed2N1-CytB-AS, complete sequence.  
ACCESSION AY640628  
VERSION AY640628.1 GI:56119185  
KEYWORDS  
SOURCE siRNA vector pSUPER-hSyn-DsRed2N1-CytB-AS  
ORGANISM siRNA vector pSUPER-hSyn-DsRed2N1-CytB-AS  
REFERENCE other sequences; artificial sequences; vectors.  
AUTHORS 1 (bases 1 to 4829)  
TITLE Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.  
JOURNAL Long-term in vivo and in vitro AAV-2-mediated RNA interference in  
rat retinal ganglion cells and cultured primary neurons  
Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)  
PUBMED 15582578  
REFERENCE 2 (bases 1 to 4829)  
AUTHORS Michel,U. and Kuegler,S.  
TITLE Direct Submission  
JOURNAL Submitted (27-MAY-2004) Neurology, University of Goettingen,  
Waldweg 33, Goettingen 37073, Germany  
FEATURES  
source  
1. .4829  
/organism="siRNA vector pSUPER-hSyn-DsRed2N1-CytB-AS"  
/mol\_type="other DNA"  
/db\_xref="taxon:297630"  
rep\_origin 135. .441  
/note="F1-Ori"  
625. .645  
promoter /note="T7 promoter"  
polyA\_signal complement(739. .899)  
/note="SV40-pA; SV40 polyadenylation sequence"  
intron complement(899. .1039)  
/note="SV40 chimeric intron derived from pCI-NEO"  
gene complement(1045. .1722)  
/gene="DsRed2"  
CDS complement(1045. .1722)  
/gene="DsRed2"  
/note="derived from Clontech's pDsRed2-N1"  
/codon\_start=1  
/transl\_table=11  
/product="DsRed2"  
/protein\_id="AAV73958.1"  
/db\_xref="GI:56119186"  
/translation="MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGGRPYEGHNTV  
KLKVTKGGLPFAWDILSPQYQYGSKVYKHPADIPDYKLSFPEGFKWERVWNVEDG  
GVAIVTQDSSLQDGCFTYKVFIGNFSPSDGPMQKTMGWEASTERLYPRDGLKGE  
THKALKLKDGGHYLVEFKSIYMAKKPVQLPGYYVDAKLDITSHNEDYTIIVEYERTE  
GRHLFL"  
misc\_feature complement(1741. .1759)  
/note="MCS; multiple cloning site"  
promoter complement(1759. .2228)  
/note="HSYN-promoter; human synapsin 1 gene promoter (-422

to +53)"  
2240. .2455  
/note="disabled human H1 RNA promoter"  
2456. .2587  
/note="Cytb-AS; partial human cytochrome B antisense  
sequence"  
complement(2640. .2658)  
/note="T3 promoter"  
3026. .3693  
/note="pUC Ori"  
complement(3841. .4701)  
/gene="Amp"  
complement(3841. .4701)  
/gene="Amp"  
/note="ampicillin resistance"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="AAV73959.1"  
/db\_xref="GI:56119187"  
/translation="MSIOHFRVALIPFAFCLPVFAHPETLVKVKDAEDQLGARVGY  
IELDINSKILDESFRPEERFPMNSTFKYLLCGAVLSRIDAGQEQLEGRIRHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGSPKEILTAFHNGDHTRL  
DRWEPELNEALPNDERDTMPVAMATTLRKLLTGLLTLASRQQLIDMEADKVAGPL  
LRSLPAGMFIADKSGAGERSRGIILALGPDGKSRIVIVITGSGATMDERRNRQILA  
EIGASLIKHW"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,11e-110 Length: 4829  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY640628 (1-4829)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGly 20  
Db 1722 ATGGCCTCCTCCGAGACGTCATCACCGAGTTCAATGCGCTTCAAGTGCGCATGAGGCC 1663  
Qy 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyArgProTyrGlyGly 40  
Db 1662 ACCGTGAACGCCACGAGCTTGAAGATCGAGGGCGAGGGCGGCCCTACGAGGCC 1603  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 1602 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTTGGGACATC 1543  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 1542 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTGACGAAGCACCCGCCGACATCCCC 1483  
Qy 81 AspTyrLysLysLeuSerPheProGlyGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 1482 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTGAG 1423  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 1422 GACGGCGGCGTGGCGGACCGACCCAGACTCTCTCCCGACGCGCCCGTGATGCAGAGAAGACC 1303  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 1362 AAGGTGAAGTTCAATCGCGGTGAACCTTCCCTCCGACGCGCCCGTGATGCAGAGAAGACC 1303  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1302 ATGGGCTGGAGGCTCTCACCGAGCGCCTGTACCCCGGACGCGCGTGCTGAAGGGCGAG 1243  
Qy 161 IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1242 ACCCAACAAGGCCCTGAAGCTGAAGAGACGCGGCCACTACCTGTGAGTTCAAGTCCATC 1183







Db 1122 ATCACTCCCAACAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 1063  
Qy 221 H18H18leuPheleu 225  
Db 1062 CACCACCTGTTCTCTG 1048  
RESULT 37  
AJ851284 5311 bp DNA circular SYN 12-NOV-2004  
LOCUS Cloning vector pRU1104 dsRed2 gene for Ds Red fluorescent protein  
DEFINITION 2.  
ACCESSION AJ851284  
VERSION AJ851284.1 GI:55724874  
KEYWORDS Ds Red fluorescent protein 2; dsRed2 gene.  
SOURCE Cloning vector pRU1104  
ORGANISM Cloning vector pRU1104  
REFERENCE 1  
AUTHORS Karunakaran, R. and Poole, P.S.  
TITLE High throughput promoter probe vectors  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 5311)  
Poole, P.S.  
TITLE Direct Submission  
JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of  
Reading, Whiteknights, Reading, RG6 6AU, UNITED KINGDOM  
FEATURES  
source 1. 5311  
/organism="Cloning vector pRU1104"  
/mol\_type="other DNA"  
/db\_xref="taxon:299179"  
/lab\_host="Escherichia coli"  
gene 326..1003  
/gene="dsRed2"  
CDS 326..1003  
/gene="dsRed2"  
/codon\_start=1  
/transl\_table=11  
/product="Ds Red fluorescent protein 2"  
/protein\_id="CA64889.1"  
/db\_xref="GI:55724875"  
/translation="MASSENVITTEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTV  
KLVKVTGGPLPFAWDILSPQFQYGSKVYKHPADLPDYKLSPEGFKWERVNFEDG  
GVATVTQDSSLQDGCFTYKYKFIQVNFPSDGPVMQKTMGWEASTERLYPRDGLKGE  
THKALKLKDGGHYLVFPRKSIYMAKKPVQLPGYVVYDAKLDTSHNEDYTIIVEQYERTE  
GRHHLPL"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.25e-110 Length: 5311  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AJ851284 (1-5311)  
Qy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 326 ATGGCTCTCCGAGAACGTCAACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 385  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 386 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC 445  
Qy 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 446 CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTGGGACATC 505  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80

Db 506 CTGTCCCCCAGTTCCAGTACGCGCTCCAAAGGTGTAGTGAAGCACCCCGCAGATCCCC 565  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPhcGlu 100  
Db 566 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAG 625  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 626 GACGGCGCGGTGGCGCAGCTGACCCAGACTCTCTCCGACAGACGGGTGCTTCACTAC 685  
Qy 121 LySValLySPhcIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 686 AAGGTGAAGTTCAATCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGAGAAGAAC 745  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLyegLyGlu 160  
Db 746 ATGGCTGGAGGCTTCACCGAGCGCCTGTACCCCGCAGCGCGCTGTGAAGGGCGAG 805  
Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
Db 806 ACCACAAGGCCCTGAAGCTGAAGAGACGGCGCCCACTACCTGTGAGTTCAAGTTCATC 865  
Qy 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 866 TACATGGCCCAAGAACCCGTGCAAGCTGCCCCGGCTACTACTACGTGAGCGCAAGTGGAC 925  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 926 ATCACTCCCAACAGAGACTACACCATCTGTGAGAGAGTACGAGCGCACCGAGGGCCGC 985  
Qy 221 H18H18leuPheleu 225  
Db 986 CACCACCTGTTCTCTG 1000  
RESULT 38  
AY613997 6423 bp DNA circular SYN 26-JUN-2004  
LOCUS Cloning vector pSRalphaneor, complete sequence.  
DEFINITION AY613997  
ACCESSION AY613997.1 GI:48995616  
KEYWORDS Cloning vector pSRalphaneor  
SOURCE Cloning vector pSRalphaneor  
ORGANISM Cloning vector pSRalphaneor  
REFERENCE 1 (bases 1 to 6423)  
AUTHORS Allard, V., de Leseleuc, L. and Denis, F.  
TITLE A family of mammalian expression vectors with different selection  
markers  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 6423)  
AUTHORS Allard, V., de Leseleuc, L. and Denis, F.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-2004) INRS-Institut Armand-Frappier, 531  
Boulevard des Prairies, Laval, Quebec H7V 1B7, Canada  
FEATURES  
source 1. 6423  
/organism="Cloning vector pSRalphaneor"  
/mol\_type="other DNA"  
/db\_xref="taxon:279783"  
misc\_feature 100..686  
/note="ECMV internal ribosomal entry site"  
misc\_feature 786..1360  
/note="ECMV internal ribosomal entry site"  
gene 1431..2108  
/gene="dsRED2"  
CDS 1431..2108  
/gene="dsRED2"  
/codon\_start=1  
/transl\_table=11  
/product="red fluorescent protein"  
/protein\_id="AAT48428.1"  
/db\_xref="GI:48995617"  
/translation="MASSENVITTEFMRFKVRMEGTVNGHEFEIEGEGRPYEGHNTV



KLKVTKGGLPFWMDILSPQFYGSKVYVKHPADIDYKKLSFPEGFKWERVMNFE  
GVAITVQDSSLQDGCFTYKVKFIGVNFPSDGPVMOKTMGWEASTERLYPRDGLKGE  
THKALKKDGGHYLVFVKSIYMAKKPVQLPGYVVYVDAKLIDITSHNEDYTTIVEQYERTE  
GRHHLFL"  
misc\_feature 2134..2266  
/note="contains SV40 polyA signal"  
gene complement(2294..3088)  
/gene="aphII"  
complement(2294..3088)  
/gene="aphII"  
/codon\_start=1  
/transl\_table=11  
/product="neomycin phosphotransferase"  
/protein\_id="AA148429.1"  
/db\_xref="GI:48995618"  
/translation="MIEQDGLHAGSPAAWVERLFGYDWAQOTIGCSDAVFRLSAQR  
PVLVKTDLSGALNELQDEARLSWLATTGVPAAVLDTVTAGRDWLLGEVPGDL  
LSSHLPAAEKVSIMADAMRRLHTLPATCPFDHQAKHRIERARTMEAGLVDDDDDE  
EHQGLAPAELEFARLKARMPDGLVVTGDACLFNIMVENGRFSGFIDCGRLGVADRY  
QDIALATRDIAELGGEMADRFVLXGIAAPDSQRIAFYRLDEFF"  
promoter complement(3456..3783)  
/note="SV40 early"  
gene 3914..4774  
/gene="b1a"  
3914..4774  
/gene="b1a"  
/codon\_start=1  
/transl\_table=11  
/product="ampicillin resistance protein"  
/protein\_id="AA148430.1"  
/db\_xref="GI:48995619"  
/translation="MSIQHFRVALIPEFAFCLPVFAHPETLVKVKDAEDQLGARVGY  
IEPLINSKILESFRPEERFPMWSTFKVLICGAVLSRVDAGQEQLGRRIHYSQNDLVE  
YSPTEKHLTDGMTVRELCSAATMSDNTAANLLTTIGSPKEELTAFIHNMGDHTRL  
DRWEPELNEAIPNDERDTTPMAPAATTLRKLTGELLTLRSQQLIDMEADKVAGPL  
LRSLAPAGWFIADKSGAGERGSRGIIAALGPDKPSRIIVITTTGSQATMDERNQIA  
EIGASLIKHW"  
rep\_origin 4961..5535  
/note="ColE1"  
5786..6408  
/note="SRalpha"  
promoter  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.57e-110 Length: 6423  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY613997 (1-6423)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1431 ATGGCCTCCTCCGAGAACGTCAATCACCGAGTTCAATGCGCTTCAAGGTCCGCAATGAGGGC 1490  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 1491 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 1550  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrPaspIle 60  
DB 1551 CACAACACCGTGAAGCTGAGACCAAGGCGCGCCCTTCCCTTCCGCTGGACATC 1610  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 1611 CTGTCCCCCAGTTCAGTACGGCTCCCAAGGTAGTGAAGCAACCCGCGCAATCCCC 1670  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 1671 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACTTCGAG 1730

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 1731 GACGCGCGCGGTGGCGACCGGTACCCAGACTCTCCCTGCAGAGACGGCTGCTCATCTAC 1790  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 1791 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCGTGATGAGAAGAACCC 1850  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 1851 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 1910  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 1911 ACCCAAGAGCCCTGAAGCTGAAGACGCGCGCACTACTGCTGAGTTCAGTTCATC 1970  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 1971 TACATGGCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGAACGCAAGCTGAC 2030  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 2031 ATCACTTCCCAACAACGAGACTACCACTCGTGAGAGACTACGACGACGACGAGCGCGC 2090  
QY 221 HisHisLeuPheLeu 225  
DB 2091 CACCACCTGTTCCTG 2105  
RESULT 39  
AY640634/c 6990 bp DNA circular SYN 01-JUL-2005  
LOCUS AY640634  
DEFINITION SiRNA vector pAAV9(5)-hSyn-DsRed2N1-EGFP-siRNA, complete sequence.  
ACCESSION AY640634  
VERSION AY640634.1 GI:56119203  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
2 (bases 1 to 6990)  
AUTHORS Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.  
TITLE Long-term in vivo and in vitro AAV-2-mediated RNA interference in  
JOURNAL rat retinal ganglion cells and cultured primary neurons  
PUBMED 15582578 Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
2 (bases 1 to 6990)  
AUTHORS Michel,U., Malik,I. and Kuegler,S.  
TITLE Direct Submision  
JOURNAL Submitted (01-JUN-2004) Neurology, University of Goettingen,  
Waldweg 33, Goettingen 37073, Germany  
FEATURES  
source  
1..6990  
/organism="SiRNA vector pAAV9(5)-hSyn-DsRed2N1-EGFP-siRNA"  
/mol\_type="other DNA"  
/db\_xref="taxon:297636"  
complement(38..193)  
/note="SV40-pA'; SV40 polyadenylation sequence"  
complement(198..341)  
/note="SV40 chimeric intron derived from pCI-NEO"  
complement(348..1025)  
/gene="DsRed2"  
/note="derived from Clontech's pDsRed2-N1"  
complement(348..1025)  
/gene="DsRed2"  
/codon\_start=1  
/transl\_table=11  
/product="DsRed2"  
/protein\_id="AAV73970.1"  
/db\_xref="GI:56119204"  
/translation="MASSENVTFEPMKFKVRMEGTVNGHEFTEGEGGRPYEGHNTV  
KLKVTKGGLPFWMDILSPQFYGSKVYVKHPADIDYKKLSFPEGFKWERVMNFE  
GVAITVQDSSLQDGCFTYKVKFIGVNFPSDGPVMOKTMGWEASTERLYPRDGLKGE  
THKALKKDGGHYLVFVKSIYMAKKPVQLPGYVVYVDAKLIDITSHNEDYTTIVEQYERTE  
GRHHLFL"



	misc_feature	complement(1044..1062)
	/note="MCS; multiple cloning site"	
promoter	complement(1062..1531)	
	/note="HSRN-promoter; human synapsin 1 gene promoter (-422 to +53) "	
promoter	1543..1758	
	/note="human H1 RNA promoter"	
misc_feature	1759..1822	
	/note="s1BGF2; 64 bp siRNA sequence"	
gene	complement(1827..4090)	
	/gene="9(5) "	
	/note="partial sequence from the non-coding porcine RNA UM 9(5) "	
repeat_region	4110..4250	
	/note="right ITR"	
gene	/rpt_family="inverted"	
	5167..6027	
	/gene="Amp"	
CDS	5167..6027	
	/gene="Amp"	
	/note="ampicillin resistance"	
	/codon_start=1	
	/transl_table=11	
	/product="beta-lactamase"	
	/protein_id="AAV73971.1"	
	/db_xref="GI:56119205"	
	/translation="MSIQFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGARVGY	
	IELDINSGKILSEFRPEERFPMSTFKYLICGAVLSRIDAGEQLGRIIHSQNDLVE	
	YSPVTEKHLTDGMTVRELCSAAITMSDNTANLLLTIGGPKELTAFHMGDHTRL	
	DRWEPELNEAI PNDERDITMPVAMATTLKLLTGSELLTASRQQLIDMEADKVAGPL	
	LRSLPAGWFIADKSGAGERSRGIIALGPDGKPSRIVITYTTSQATMDERNQIA	
	EIGASLTKHW"	
	6848..6988	
repeat_region	/note="left ITR"	
	/rpt_family="inverted"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.74e-110	Length: 6990
Score:	1186.00	Matches: 219
Percent Similarity:	98.22%	Conservative: 2
Best Local Similarity:	97.33%	Mismatches: 4
Query Match:	97.69%	Indels: 0
DB:	11	Gaps: 0
US-10-006-922A-12 (1-225) x AY640634 (1-6990)		
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20	
DB	1025 ATGCCCTCCTCCGAGAACGTCAACCCGAGTTCATGCCCTTCAAGGTGCGCATGAGGGC 966	
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40	
DB	965 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTTACGAGGGC 906	
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60	
DB	905 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGGACATC 846	
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80	
DB	845 CTGTCCCCCAGTTCCAGTACGAGCTCCAAAGGTATCGTGAAGCACCCCGCCGACATCCC 786	
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100	
DB	785 GACTACAAGAGCTGTCTTCCCTCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAG 726	
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120	
DB	725 GACGGCGGCGTGGCGAACC GTGACCCAGGACTCTCTCCGACGAGGACGGCTGCTTCACTAC 666	
QY	121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140	

Db	665	AAGGTGAAGTTTCATCGGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGGAAGAAC	606
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	605	ATGGGCTGGAGGCGCTCCACCGAGCGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAG	546
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	545	ACCCACAAGGCCCTGAAGCTGAAGACGCGCGCCACTACTGTGTGAAGTTCAAGTCCATC	486
Qy	161	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAlaSerLysLeuAsp	200
Db	485	TACATGGCCAGAAGCCCGTGCACTGCGCCGCTACTACTACTGAGACGCCAAGCTGAC	426
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg	220
Db	425	ATCACCTCCCAACAAGAGACTACACCATCGTGGAGCAGTACGAGCGACCGAGGCGCGC	366
Qy	221	HisHisLeuPheLeu	225
Db	365	CACCACCTGTTCTCTG	351
RESULT 40			
AY640633/c			
LOCUS	AY640633	7058 bp	DNA
DEFINITION	siRNA vector pAAV9(5)-hSyn-DsRed2N1-CytB-AS-ohneNot, complete		
ACCESSION	AY640633		
VERSION	AY640633.1	GI:56119200	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 (bases 1 to 7058)		
TITLE	Michel,U., Malik,I., Ebert,S., Bahr,M. and Kugler,S.		
JOURNAL	Long-term in vivo and in vitro AAV-2-mediated RNA interference in		
PUBMED	rat retinal ganglion cells and cultured primary neurons		
REFERENCE	Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)		
AUTHORS	2 (bases 1 to 7058)		
TITLE	Michel,U., Malik,I. and Kuegler,S.		
JOURNAL	Direct Submission		
PUBMED	Submitted (01-JUN-2004) Neurology, University of Goettingen,		
REFERENCE	Waldweg 33, Goettingen 37073, Germany		
AUTHORS	Location/Qualifiers		
TITLE	1. 7058		
JOURNAL	/organism="siRNA vector		
PUBMED	pAAV9(5)-hSyn-DsRed2N1-CytB-AS-ohneNot"		
REFERENCE	/mol_type="other DNA"		
AUTHORS	/db_xref="taxon:297635"		
TITLE	1. 7141		
JOURNAL	/note="left ITR"		
PUBMED	/rpt_family="inverted"		
REFERENCE	complement(38..193)		
AUTHORS	/note="SV40-pA'; SV40 polyadenylation sequence"		
TITLE	complement(198..341)		
JOURNAL	/note="SV40 chimeric intron derived from pCI-NEO"		
PUBMED	complement(491..1168)		
REFERENCE	/gene="DsRed2"		
AUTHORS	complement(491..1168)		
TITLE	/gene="DsRed2"		
JOURNAL	/note="derived from Clontech's pDsRed2-N1"		
PUBMED	/codon_start=1		
REFERENCE	/transl_table=11		
AUTHORS	/product="DsRed2"		
TITLE	/protein_id="AAV73969.1"		
JOURNAL	/db_xref="GI:56119202"		
PUBMED	/translation="MASSENVITTEMRFKVRMEGTVNGHFEIEGEGRPYEAGNTV		
REFERENCE	KLKVTGGPLPEAMDILSPQFGSKVYVGPADIPDKLSFPEGFKVRVMNFEDGKGE		
AUTHORS	GVAIVTQDSSLQDGCFTYKVKFIVGNFPSPDGPVMQKTMGEASTERLYPRDVLKGE		
TITLE	THKALKLKDGSHYLVFEKSIYMAKKEPVQLPGYYVVDALDITSHNEDYTIIVEQYERTE		
JOURNAL	GRHHLFL"		
PUBMED	complement(1187..1205)		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			



/note="MCS; multiple cloning site"  
complement(1205..1674)  
/note="HSYN-promoter; human synapsin 1 gene promoter (-422 to +53)"  
1686..1901  
/note="disabled human H1 RNA promoter"  
1902..2033  
/note="Cytb-AS; partial human cytochrome B antisense sequence"  
complement(2038..4301)  
/gene="9(5)"  
/note="partial sequence from the non-coding porcine RNA UM 9(5)"  
4321..4461  
/note="right ITR"  
/rpt\_family="inverted"  
5378..6238  
/gene="Amp"  
5378..6238  
/gene="Amp"  
/note="ampicillin resistance"  
/codon\_start=1  
/transl\_table=1  
/product="beta-lactamase"  
/protein\_id="AAV73968.1"  
/db\_xref="GI:56119201"  
/translation="MSIQHFRVALIPFAAFCLPVPAHPELTVKVKDAEDDLGARVGY  
IELDINSGKILESFRPEERFPMWSTFKVLGCAVLSRIDAQEQELRRIRHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAATMSDNTAANLLTTGGPKELTAFIHNMGDHTRL  
DRMEPELNEAI PNDERDITMVPAMATTLRLTLGELTLASRQQLIDMEADKYAGPL  
LRSAIPAGWFIADKSGAGERSGRIIALGPDGKPSRIIVITYTGSQATMDERNQIA  
ELGASLIKHW"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.76e-110 Length: 7058  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AY640633 (1-7058)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1168 ATGGCCTCCTCCGAGAAAGTCATCACCGAGTTCAATGCCCTTCAAGGTGGCATGAGGGC 1109  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 1108 ACCGTGAACGGCCAGAGTTGAGATCGAGGCGGAGGGCGGCCCTTACGAGGGC 1049  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 1048 CACAACACCGTGAAGCTGAAGTGACCAAGGCGGCCCTGCTCCCTTGGGACATC 989  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 988 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGACGACACCCCGGCATATCCCC 929  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 928 GACTACAGAGAGCTGCTCTTCCCCGAGGCGCTTCAAGTGGAGCGCGTGAATCTCGAG 869  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 868 GACGGCGGCGTGGCGGACCGTGAACCAAGACTCTCTCCCTGCAAGACGGCTTCAATCTAC 809  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 808 AAGGTGAAGTTCAATCGGCTGAACCTTCCCTCCGACGGCCCGTGTATGCAGAGAAGACC 749  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

DB 748 ATGGCTGGAGGCTTCCACGAGCGCCTGTACCCCCCGACGGCGTGTGAAGGGCAG 689  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 688 ACCACAAAGGCCCTGAAGCTGAAGAGACGGCGGCCACTTACTGTGAGTTCAAGTTCATC 629  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 628 TACATGGCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACGCCAAGCTGCAC 569  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 568 ATCACTCCACAAACGAGACTACACCATCGTGAGCAGTACGACGACCGACCGGCCGC 509  
QY 221 HisHisLeuPheLeu 225  
DB 508 CACCACCTGTTCTTG 494

RESUL'T 41  
AY640630/c 7147 bp DNA circular SYN 01-JUL-2005  
LOCUS  
DEFINITION  
SIRNA vector PAAV9(5) -CMV-DsRed2N1-CytB-AS-ohneNot, complete  
ACCESSION  
AY640630  
AY640630.1 GI:56119191  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

SIRNA vector PAAV9(5) -CMV-DsRed2N1-CytB-AS-ohneNot  
SIRNA vector PAAV9(5) -CMV-DsRed2N1-CytB-AS-ohneNot  
other sequences; artificial sequences; vectors.

REFERENCE  
1 (bases 1 to 7147)  
AUTHORS  
TITLE  
Long-term in vivo and in vitro AAV-2-mediated RNA interference in  
rat retinal ganglion cells and cultured primary neurons  
JOURNAL  
Biochem. Biophys. Res. Commun. 326 (2), 307-312 (2005)  
PUBMED  
15582578  
REFERENCES  
2 (bases 1 to 7147)  
AUTHORS  
TITLE  
Direct Submission  
Submitted (01-JUN-2004) Neurology, University of Goettingen,  
Waldweg 33, Goettingen 37073, Germany  
FEATURES  
Location/Qualifiers

source

/organism="SIRNA vector  
PAAV9(5) -CMV-DsRed2N1-CytB-AS-ohneNot"

/mol\_type="other DNA"

/db\_xref="taxon:297632"

1..141

repeat\_region

/note="left ITR"

/rpt\_family="inverted"

complement(181..338)

polyA\_signal

/note="SV40-pA'; SV40 polyadenylation sequence"

complement(341..481)

intron

/note="SV40 chimeric intron derived from pCI-NEO"

complement(491..1168)

gene

/gene="DsRed2"

complement(491..1168)

CDS

/gene="DsRed2"

/note="derived from Clontech's pDsRed2-N1"

/codon\_start=1

/transl\_table=11

/product="DsRed2"

/protein\_id="AAV73963.1"

/db\_xref="GI:56119193"

/translation="MASSENVITTEFMRFKVRMEGTVNGHEFIEGEGRPYEGHNTV  
KLKVTGKGPLPAMWILSPQFGYSKYVVKHPADIPYKKLSFPEGFKMERVMNFEIDG  
GVAITYQDSLQDGFYIKYKFTGVNFPSPDGPVMOKTMGEASTERLYPRDGVLTGE  
THKALKLKDGGHYLVFEFKSIYNAKKPVQLPGYYVDAKLDTSHNEDYTIIVEQYERTE  
GRHHLFL"

misc\_feature

complement(1187..1206)

/note="MCS'; multiple cloning site"

promoter

complement(1217..1739)

/note="MCMV promoter; mcmv promoter (-491 to +36)"



promoter 1775..1990  
/note="disabled human H1 RNA promoter"  
misc\_feature 1991..2122  
/note="Cytb-AS; partial human cytochrome B antisense  
sequence"  
gene complement(2127..4390)  
/gene="9(5)"  
/note="partial sequence from the non-coding porcine RNA UM  
9(5)"  
repeat\_region 4410..4550  
/note="right ITR"  
/rpt\_family="inverted"  
gene 5467..6327  
/gene="Amp"  
CDS 5467..6327  
/gene="Amp"  
/note="ampicillin resistance"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="AAV3962.1"  
/db\_xref="GI:56119192"  
/translation="MSIOHFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGARVGY  
IELDLSGKILESFRPERFPMSTPKVLLCGAVISRIDAGQQLGRIRIYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGSPELTAFLHNMGDHVTRL  
DRWEPELNEAIPNDERDTTMPVAWATTLRKLTIGLILLTASRQQLDWMEDKVAGPL  
LRSAIPAGWFIADKSGAGRGSRGIIAALGPDKSRIVIVITTSQATMDERNRQIA  
EIGASLIKHW"

ORIGIN

Alignment Scores:  
Pred. No.: 1.79e-110 Length: 7147  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AY640630 (1-7147)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1168 ATGGCTCTCCGAGAACGTCAACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 1109  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 1108 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC 1049  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 1048 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCGCTTGGGACATC 989  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 988 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCACATCCC 929  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 928 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 869  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 868 GACGCGCGCGTGGCGACCGTGAACCTCCTCCCTGACGAGCGGCTGTTCATCTAC 809  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 808 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAAGACC 749  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 748 ATGGGCTGGAGGCTCCACCGAGGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 689  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180

Db 688 ACCCAAGAGGCCCTGAAGCTGAAGGACGGCGGCCCACTACCTGTGTGAGTTCAGTTCATC 629  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 628 TACATGGCCAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCGCCAAGCTGGAC 569  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 568 ATCACTCCACAAACGAGACTACACCATCGTGGAGCACTACGAGCGCACCGAGGCGCC 509  
QY 221 HisHisLeuPheLeu 225  
Db 508 CACCACCTGTTCCTG 494

RESULT 42  
LOCUS CQ849511 7616 bp DNA linear PAT 23-AUG-2004  
DEFINITION Sequence 8 from Patent WO2004067751.  
ACCESSION CQ849511  
VERSION CQ849511.1 GI:51507515  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Couillard-Despres,S., Karl,C., Kuhn,H.G. and Aigner,L.  
TITLE Use of regulatory sequences for specific, transient expression  
JOURNAL Patent: WO 2004067751-A 8 12-AUG-2004;  
Klinikum der Universitaet Regensburg (DE)  
FEATURES  
source location/Qualifiers  
1..7616  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="fusion construct homo sapiens and DsRed 2"

ORIGIN

Alignment Scores:  
Pred. No.: 1.93e-110 Length: 7616  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CQ849511 (1-7616)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 3606 ATGGCTCTCCGAGAACGTCAACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 3665  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 3666 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 3725  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 3726 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTCGCTTGGGACATC 3785  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 3786 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGACATCCCC 3845  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 3846 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 3905  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 3906 GACGGCGCGTGGCGACCGTGAACCAAGACTCTCCTCGACGAGCGGCTGTTCATCTAC 3965







Qy	41	H1sAsnThrValIysLeuLysValThrIysGlyProLeuProPheAlaTrpAspIle	60
Db	8607	CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGCGCTGGACATC	8548
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	8547	CTGTCCCCCAGTTCACAGTACGGCTCCAAGGTGTACGTGAAGCACCCCCGCACATCCCC	8488
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	8487	GACTACAGAAGCTGTCTCTTCCCCAGGGCTTCAAGTGGAGCCGCTGATGAACCTTCAG	8428
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	8427	GACGGCGCGTGGCGACCGTGACCAAGACTCTCTCCCTGCAGACGGCTGCTTCATCTAC	8368
Qy	121	LysValLysPheIleGlyValAsnAspProSerAspGlyProValMetGlnLysLysThr	140
Db	8367	AAGGTGAAGTTCATCGCGTGAACCTTCCCTCCGACGGCCCGTGATGCAGAAGAAGACC	8308
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	8307	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG	8248
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	8247	ACCCACAAGCCCTGAAGCTGAAGGACGGCGCCACTACCTGTGGAGTTCAAGTCCATC	8188
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	8187	TACATGGCCAGAAGCCCGTGCACTGCGCGCTACTACTACTAGTGAACGCCAAGCTGGAC	8128
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	8127	ATCACTTCCCAACAAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCGCG	8068
Qy	221	HisHisLeuPheLeu	225
Db	8067	CACCACTGTTCCTG	8053

RESULT 45  
 CS018254  
 LOCUS CS018254 10522 bp DNA linear PAT 23-FEB-2005  
 DEFINITION Sequence 21 from Patent WO2005012534.  
 ACCESSION CS018254  
 VERSION CS018254.1 GI:60220035  
 KEYWORDS  
 SOURCE .  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Alphey, L.  
 TITLE Expression systems for insect pest control  
 JOURNAL Patent: WO 2005012534-A 21 10-FEB-2005;  
 Oxitec Limited (GB)  
 FEATURES  
 source  
 1..10522  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="pLA1124"

ORIGIN

Alignment Scores:  
 Pred. No.: 2.87e-110 Length: 10522  
 Score: 1186.00 Matches: 219  
 Percent Similarity: 98.22% Conservative: 2  
 Best Local Similarity: 97.33% Mismatches: 4  
 Query Match: 97.69% Indels: 0  
 DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018254 (1-10522)

Qy	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyGly	20
Db	5028	ATGGCTCTCCGAGAAAGTCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGCG	5087
Qy	21	ThrValAsnGlyHisGluPheGluIleGlyGlyGluGlyArgProTyrGlyGly	40
Db	5088	ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGCG	5147
Qy	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	5148	CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGACATC	5207
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	5208	CTGTCCCCCAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCGCGACATCCCC	5267
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	5268	GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG	5327
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	5328	GACGGCGCGTGGCGACCGGTGACCCAGACTCTCTCCCTGCAGACGGCTGCTTCATCTAC	5387
Qy	121	LysValLysPheIleGlyValAsnAspProSerAspGlyProValMetGlnLysLysThr	140
Db	5388	AAGGTGAAGTTCATCGCGCTGAACCTTCCCTCCGACGGCCCGTGATGCAGAAGAAGACC	5447
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	5448	ATGGGCTGGGAGGCTTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG	5507
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	5508	ACCCACAAGCCCTGAAGCTGAAGGACGGCGCCACTACCTGTGAGTTCAAGTCCATC	5567
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	5568	TACATGGCCAGAAGCCCGTGCACTGCGCGCTACTACTACTAGTGAACGCCAAGCTGGAC	5627
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	5628	ATCACTTCCCAACAAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCGCG	5687
Qy	221	HisHisLeuPheLeu	225
Db	5688	CACCACTGTTCCTG	5702

RESULT 46  
 CS018256/c  
 LOCUS CS018256 10786 bp DNA linear PAT 23-FEB-2005  
 DEFINITION Sequence 23 from Patent WO2005012534.  
 ACCESSION CS018256  
 VERSION CS018256.1 GI:60220037  
 KEYWORDS  
 SOURCE .  
 ORGANISM synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE 1  
 AUTHORS Alphey, L.  
 TITLE Expression systems for insect pest control  
 JOURNAL Patent: WO 2005012534-A 23 10-FEB-2005;  
 Oxitec Limited (GB)  
 FEATURES  
 source  
 1..10786  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="pLA670"

ORIGIN

Alignment Scores:  
 Pred. No.: 2.95e-110 Length: 10786



Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018256 (1-10786)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 8727 ATGGCCTCTCCGAGAAGCTCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 8668  
QY 21 ThrValAenGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 8667 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 8608  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 8607 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGGGACATC 8548  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 8547 CTGTCCCCCAGTTCAGTACGCGCTCCAAAGGTGTACGTGAAGCACCCCGCATCCCC 8488  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 8487 GACTACAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTGAACTTCGAG 8428  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 8427 GACGCGCGCGTGGGACCGGTGACCCAGACTCTCCCTGCAAGACGCGCTTCATCTAC 8368  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 8367 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAGACC 8308  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 8307 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGGCGAG 8248  
QY 161 IleHisLySAleuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 8247 ACCCACAAGGCCCTGAAGCTGAAGACGCGGCGCACTAAGTGAAGTTCATC 8188  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 8187 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGAAGCCAACTGGAC 8128  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 8127 ATCACTCCCAACAAGAGACTACACCATGTGAGCAGTACGAGCGACCGAGGGCGGC 8068  
QY 221 HisHisLeuPheLeu 225  
DB 8067 CACCACCTGTTCTGTG 8053

RESULT 47  
CS018251/c 11251 bp DNA linear PAT 23-FEB-2005  
LOCUS Sequence 18 from Patent WO2005012534.  
DEFINITION CS018251  
ACCESSION CS018251  
VERSION CS018251.1 GI:60220032  
KEYWORDS  
SOURCE . synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 other sequences; artificial sequences.  
AUTHORS Alphaey,L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 18 10-FEB-2005;  
Oxitec Limited (GB)  
FEATURES  
Source Location/Qualifiers  
1..11251

/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA656"

ORIGIN

Alignment Scores: 3.11e-110 Length: 11251  
Pred. No.: 1186.00 Matches: 219  
Score: 98.22% Conservative: 2  
Percent Similarity: 98.22% Mismatches: 4  
Best Local Similarity: 97.33% Indels: 0  
Query Match: 97.69% Gaps: 0  
DB: 6

US-10-006-922A-12 (1-225) x CS018251 (1-11251)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 9191 ATGGCCTCTCCGAGAAGCTCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 9132  
QY 21 ThrValAenGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 9131 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 9072  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 9071 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGGGACATC 9012  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 9011 CTGTCCCCCAGTTCAGTACGCGCTCCAAAGGTGTACGTGAAGCACCCCGCATCCCC 8952  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 8951 GACTACAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTGAACTTCGAG 8892  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 8891 GACGCGCGCGTGGGACCGGTGACCCAGACTCTCCCTGCAAGACGCGCTTCATCTAC 8832  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 8831 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAGACC 8772  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 8771 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGGCGAG 8712  
QY 161 IleHisLySAleuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 8711 ACCCACAAGGCCCTGAAGCTGAAGACGCGGCGCACTAAGTGAAGTTCATC 8652  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 8651 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGAAGCCAACTGGAC 8592  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 8591 ATCACTCCCAACAAGAGACTACACCATGTGAGCAGTACGAGCGCACCGAGGGCGGC 8532  
QY 221 HisHisLeuPheLeu 225  
DB 8531 CACCACCTGTTCTGTG 8517

RESULT 48  
CS018250/c 11570 bp DNA linear PAT 23-FEB-2005  
LOCUS Sequence 17 from Patent WO2005012534.  
DEFINITION CS018250  
ACCESSION CS018250  
VERSION CS018250.1 GI:60220031  
KEYWORDS  
SOURCE . synthetic construct  
ORGANISM synthetic construct



other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alphaey,L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 17 10-FEB-2005;  
Oxitec Limited (GB)

FEATURES  
source 1. .11570  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA517"

ORIGIN

Alignment Scores:  
Pred. No.: 3.22e-110 Length: 11570  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018250 (1-11570)

QY 1 MetArgSerSerLySAsnValIlelySGluPheMetArgPheLySValArgMetGluGly 20  
DB 8886 ATGGCCTCTCCGAGAACGTATCACCAGTTTCATGCGCTTCAAGTGGCATGAGGGC 8827  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 8826 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 8767

QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 8766 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTGGGACATC 8707  
QY 61 LeuSerProGlnPheGlnTyrglySerLySValTyrgValLyHisProAlaAspIlePro 80  
DB 8706 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGACATCCCC 8647

QY 81 AspTyrLySlySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
DB 8646 GACTACAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 8587  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySphenIleTy 120  
DB 8586 GACGGCGCGGTGGCGACCGTGACCCAGACTCTCTCCCTGCAGACGGCTGCTCATCTAC 8527

QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
DB 8526 AAGGTGAAGTTTCATCGCGTGAACCTTCCCTCCAGCGGCCGTGATGCAGAAAGAGACC 8467  
QY 141 MetGlyTrrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLySGlyGlu 160  
DB 8466 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 8407

QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySertile 180  
DB 8406 ACCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGGTGAAGTTCAAGTCCATC 8347  
QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrrTyrrValAspSerLySLeuAsp 200  
DB 8346 TACATGGCCAAAGAGCCCGTGCAAGTCCCGGCTACTACTACGTGGACGCCAAGCTGGAC 8287

QY 201 IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrgLuarGThrGluGlyArg 220  
DB 8286 ATCACTTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCG 8227

QY 221 HisHisLeuPheLeu 225  
DB 8226 CACCACTGTTCCTG 8212

RESULT 49

CS018255  
LOCUS CS018255 11867 bp DNA linear PAT 23-FEB-2005  
DEFINITION Sequence 22 from Patent WO2005012534.  
ACCESSION CS018255  
VERSION CS018255.1 GI:60220036  
KEYWORDS  
SOURCE  
ORGANISM  
other sequences; artificial sequences.

REFERENCE 1  
AUTHORS Alphaey,L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 22 10-FEB-2005;  
Oxitec Limited (GB)

FEATURES  
source 1. .11867  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA1188"

ORIGIN

Alignment Scores:  
Pred. No.: 3.32e-110 Length: 11867  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CS018255 (1-11867)

QY 1 MetArgSerSerLySAsnValIlelySGluPheMetArgPheLySValArgMetGluGly 20  
DB 6373 ATGGCCTCTCCGAGAACGTATCACCAGTTTCATGCGCTTCAAGTGGCATGAGGGC 6432  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgluGly 40  
DB 6433 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 6492

QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 6493 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTGGGACATC 6552  
QY 61 LeuSerProGlnPheGlnTyrglySerLySValTyrgValLyHisProAlaAspIlePro 80  
DB 6553 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGACATCCCC 6612

QY 81 AspTyrLySlySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
DB 6613 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 6672  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySphenIleTy 120  
DB 6673 GACGGCGCGTGCGACCGTGACCCAGACTCTCTCTGCAGACGGCTGCTCATCTAC 6732

QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
DB 6733 AAGGTGAAGTTTCATCGGGTGAACCTTCCCTCCGACGGCCCGTGATGCAGAAAGAGACC 6792  
QY 141 MetGlyTrrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLySGlyGlu 160  
DB 6793 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGACGGCGTGTGAAGGGCGAG 6852

QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySertile 180  
DB 6853 ACCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACTGTGAGTTCAAGTCCATC 6912

QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrrTyrrValAspSerLySLeuAsp 200  
DB 6913 TACATGGCCAAAGCCCGTGCAAGTCCCGGCTACTACTACGTGAGCGCAAGCTGGAC 6972

QY 201 IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrgLuarGThrGluGlyArg 220



Db 6973 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAGTACGACGACCGAGGCGCG 7032

Qy 221 HSHISLeuPheLeu 225  
|||||  
7033 CACCACCTGTTCTCTG 7047

RESULT 50  
CS018249 11920 bp DNA linear PAT 23-FEB-2005  
LOCUS Sequence 16 from Patent WO2005012534.  
DEFINITION CS018249  
ACCESSION CS018249  
VERSION CS018249.1 GI:60220030  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 16 10-FEB-2005;  
Oxitec Limited (GB)  
FEATURES  
source 1. .11920  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA513"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.34e-110 Length: 11920  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: Gaps: 6  
US-10-006-922A-12 (1-225) x CS018249 (1-11920)  
Qy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
|||||  
Db 7953 ATGGCCTCCTCCGAGAACGTATCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 8012  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||||  
Db 8013 ACCGTGAACGGCCACGAGTTCCGAGATCGAGGGCGGAGGGCGCCCTACGAGGGC 8072  
Qy 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||  
Db 8073 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGCGCCCTGCTCCCTTCCGCTGGGACATC 8132  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
|||||  
Db 8133 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGACGTGAAGCACCCCGCCGACATCCCC 8192  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
|||||  
Db 8193 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAATGACTTCCAG 8252  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||  
Db 8253 GACGGCGGCGGTGGCGGACCGGTGACCCAGACTCTCCCTGCAAGGAGCGGCTGATCTTAC 8312  
Qy 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
|||||  
Db 8313 AAGGTGAAGTTCATCGGCGGTGAATCTTCCCTCCGACGGGCCCTGATGCAAGAAAGACC 8372  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
|||||  
Db 8373 ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCGGAGCGGCGTGTGAAGGGCGAG 8432  
Qy 161 IleHisLySAlaLeuLyLeuLySAspGlyGlnHisTyrLeuValGluPheLySLeuThr 180

Db 8433 ACCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTGTGTGAGTTCAAGTTCATC 8492

Qy 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
|||||  
Db 8493 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTGAGCGCAAGCTTGAC 8552

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||||  
Db 8553 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAGTACGAGCGACCGAGGCGCGC 8612

Qy 221 HSHISLeuPheLeu 225  
|||||  
Db 8613 CACCACCTGTTCTCTG 8627

RESULT 51  
CS018257/c 14720 bp DNA linear PAT 23-FEB-2005  
LOCUS Sequence 24 from Patent WO2005012534.  
DEFINITION CS018257  
ACCESSION CS018257  
VERSION CS018257.1 GI:60220038  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Expression systems for insect pest control  
JOURNAL Patent: WO 2005012534-A 24 10-FEB-2005;  
Oxitec Limited (GB)  
FEATURES  
source 1. .14720  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pLA1038"  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.32e-110 Length: 14720  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: Gaps: 6  
US-10-006-922A-12 (1-225) x CS018257 (1-14720)  
Qy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
|||||  
Db 7382 ATGGCCTCCTCCGAGAACGTATCATCACCGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 7323  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
|||||  
Db 7322 ACCGTGAACGGCCACGAGTTCCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC 7263  
Qy 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
|||||  
Db 7262 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTCGCTGGGACATC 7203  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
|||||  
Db 7202 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGACGTGAAGCACCCCGCCGACATCCCC 7143  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
|||||  
Db 7142 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAATGAACTTCCAG 7083  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||||  
Db 7082 GACGGCGGCGGTGGCGGACCGGTGACCCAGACTCTCCCTGCAAGGAGCGGCTGCTCATCTTAC 7023  
Qy 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140



Db 7022 AAGGTGAAGTTCATCGGCGGTGAACTTCCCTCCGACGGCCCCCTGATGCAGAGAAGACC 6963

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 6962 ATGGGCTGGAGGCGCTCCACCGACGAGCGCTGTACCCCCCGGACGGCGTGTGAAGGGCGAG 6903

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 6902 ACCCAAGAGCCCTGAAGCTGAAGAGCGGCGGCACTACCTGGTGAAGTTCAAGTCCATC 6843

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 6842 TACATGGCCAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGAACGCCAAGCTGAC 6783

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 6782 ATCACTCCCAACAAGGACTACCACTCGTGAGCAAGTACGAGCGCACCGAGGGCGGC 6723

Qy 221 HisHisLeuPheLeu 225

Db 6722 CACCACCTGTTCTCG 6708

RESULT 52  
CQ981073/c  
LOCUS CQ981073 16157 bp DNA linear PAT 25-JAN-2005  
DEFINITION Sequence 3 from Patent WO2005003364.  
ACCESSION CQ981073  
VERSION CQ981073.1 GI:58190337  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Stable integrands  
JOURNAL Patent: WO 2005003364-A 3 13-JAN-2005;  
Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1. .16157  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="predicted sequence of pLA1025 construct"

ORIGIN

Alignment Scores:  
Pred. No.: 4.83e-110 Length: 16157  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CQ981073 (1-16157)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 14524 ATGGCCTCCTCCGAGAAGCTCATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG 14465

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 14464 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCCCTACGAGGGC 14405

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60

Db 14404 CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCTTGCCTGGGACATC 14345

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 14344 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTGTACGTGAAGCAACCCCGCAGACATCCCC 14285

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

Db 14284 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG 14225

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 14224 GACGGCGCGTGGCGGACCGGTGACCCAGGACTCTCTCCCTGCAGAGACGGCTGCTTCTAC 14165

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 14164 AAGGTGAAGTTTCATGGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAGAAGACC 14105

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 14104 ATGGGCTGGAGGCTTCCACCGAGCGCCTGTACCCCCCGGACGGCGTGTGAAGGGCGAG 14045

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 14044 ACCCAAGAGCCCTGAAGCTGAAGAGCGGCGGCACTACCTGGTGAAGTTCAGTCCATC 13985

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 13984 TACATGGCCAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCGCCAAGCTGAC 13925

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 13924 ATCACTCCCAACAAGGACTACCACTCGTGAGCAAGTACGAGCGCACCGAGGGCGGC 13865

Qy 221 HisHisLeuPheLeu 225

Db 13864 CACCACCTGTTCTCG 13850

RESULT 53  
CQ981074/c  
LOCUS CQ981074 16157 bp DNA linear PAT 25-JAN-2005  
DEFINITION Sequence 4 from Patent WO2005003364.  
ACCESSION CQ981074  
VERSION CQ981074.1 GI:58190338  
KEYWORDS  
SOURCE .  
ORGANISM synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Alphey, L.  
TITLE Stable integrands  
JOURNAL Patent: WO 2005003364-A 4 13-JAN-2005;  
Oxitec Limited (GB)  
FEATURES  
source location/Qualifiers  
1. .16157  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="predicted sequence of pLA1125 construct"

ORIGIN

Alignment Scores:  
Pred. No.: 4.83e-110 Length: 16157  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x CQ981074 (1-16157)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 14524 ATGGCCTCCTCCGAGAAGCTCATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG 14465

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 14464 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCCCTTACGAGGGC 14405

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60



Db 14404 CACAAACCCTGAAGCTGAAGGTGACCAAGGGCGGGCCCCCTCGCCTTGCGACATC 14345

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 14344 CTGTCCCCCAAGTTCAGTACGGCTCCAGGTGTACGTGAAGCACCCCCGACATCCCC 14285

Qy 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100

Db 14284 GACTACAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 14225

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 14224 GACGGCGCGTGGCGACCGGTGACCCAGACTCTCCCTGCAGAGCGCGCTTCATCTAC 14165

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 14164 AAGGTGAAGTTCATCGGCGGTGAACCTCCCTCCGACGGCCCCGTGATGCAAGAAGACC 14105

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 14104 ATGGCTGGAGGCTCTCACCGAGCGCTGTACCCCCCGACGGCGGTGTGAAGGGCGAG 14045

Qy 161 IleHisLysValLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 14044 ACCCACAAGGCCCTGAAGCTGAAGAGCGGCGCACTACCTGTGTGAATTCAAGTCCATC 13985

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 13984 TACATGGCCAAAGAGCCCGTGCAAGTGCAGCGCGCTACTACTACGTGACGCCAAAGCTGAC 13925

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg 220

Db 13924 ATCACTCCCAACAAGAGACTACACCATCGTGAAGCAAGTACGAGCGACCGAGGGCGGC 13865

Qy 221 HisHisLeuPheLeu 225

Db 13864 CACCACCTGTCTCTG 13850

RESULT 54

AX686888 LOCUS AX686888 898 bp DNA linear PAT 31-MAR-2003

DEFINITION Sequence 11 from Patent WO0127150.

ACCESSION AX686888

VERSION AX686888.1 GI:29409468

KEYWORDS

SOURCE Discosoma sp.

ORGANISM Discosoma sp.

REFERENCE 1 Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.

AUTHORS Lukyanov,S.A., Fradkov,A.F., Labas,Y.A., Matz,M.V. and Teresikh,A.

TITLE Anthozoa derived chromo/fluoroproteins and methods for using the same

JOURNAL Patent: WO 0127150-A 11 19-APR-2001;

FEATURES SOURCE Clontech Laboratories Inc. (US)

1. .898 Location/Qualifiers

/organism="Discosoma sp."

/mol\_type="unassigned DNA"

/db\_xref="taxon:86600"

ORIGIN

Alignment Scores:

Pred. No.: 2.28e-111 Length: 898

Score: 1184.00 Matches: 219

Percent Similarity: 98.22% Conservative: 2

Best Local Similarity: 97.33% Mismatches: 4

Query Match: 97.53% Indels: 0

DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x AX686888 (1-898)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyGly 20

Db 93 ATGAGCTCTTCCAGAAGATGTATCAAGAGGTTTCATGAGGTTTAAAGTTTCGATGAAGGA 152

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 153 ACGGTCAATGGGCACGAGTTTGAATAAGAAGCGGAAGAGAGGGGAGGCCATACGAAGGC 212

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

Db 213 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTTGCCATTGCTTGGGATATT 272

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80

Db 273 TTGTCAACCAATTTCAGTATGAAGCAAGTATATGTCAAGCACTTGCCGACATACCA 332

Qy 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100

Db 333 GACTATPAAAAAGCTGTCAATTCTCGAAGATTAAATGGGAAGGTCATGAACCTTGAA 392

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 393 GACGTGGCGTCTTACTGTAAACCAAGATTCCAGTTTGCAGATGGCTGTTTCATCTAC 452

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

Db 453 AAGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTGTAACAAAAGAAACA 512

Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 513 ATGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAGAGAGAG 572

Qy 161 IleHisValLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180

Db 573 ATTCAATAAGGCTCTGAAGCTGAAGAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 632

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 633 TACATGGCAAGAAGCCCTGTGACGCTACCAAGGTACTACTATGTTGACTCCAAACTGAT 692

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg 220

Db 693 ATAACAAGCCCAACAAGAACTATACATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC 752

Qy 221 HisHisLeuPheLeu 225

Db 753 CACCATCTGTCTCTT 767

RESULT 55

AY916793 LOCUS AY916793 3441 bp DNA circular SYN 30-MAR-2005

DEFINITION Dual fluorescent protein cloning vector pGRFP, complete sequence.

ACCESSION AY916793

VERSION AY916793.1 GI:60101687

KEYWORDS

SOURCE Dual fluorescent protein cloning vector pGRFP

ORGANISM Dual fluorescent protein cloning vector pGRFP

REFERENCE 1 other sequences; artificial sequences; vectors.

AUTHORS 1 (bases 1 to 3441)

TITLE Choe,J., Guo,H.H. and van den Engh,G.

JOURNAL A dual-fluorescence reporter system for high-throughput clone

PUBMED characterization and selection by cell sorting

15767274 (er) Nucleic Acids Res. 33 (5), E49 (2005)

REFERENCE 2 (bases 1 to 3441)

AUTHORS Choe,J., Guo,H.H. and van den Engh,G.

TITLE Direct Submission

JOURNAL Submitted (01-FEB-2005) Laboratory of Ger van den Engh, Institute

for Systems Biology, 1441 N 34th St, Seattle, WA 98103-8904, USA

FEATURES SOURCE Location/Qualifiers

1. .3441

/organism="Dual fluorescent protein cloning vector pGRFP"

/mol\_type="other DNA"



/db\_xref="taxon:314306"  
140\_.1630  
/gene="GRFP"  
140\_.1630  
CDS  
/gene="GRFP"  
/note="fusion protein containing GFPmut3.1 and DsRed-T3 with an intervening linker region"  
/codon\_start=1  
/transl\_table=11  
/product="GFP-DsRed fusion protein"  
/protein\_id="AA13969.1"  
/db\_xref="GI:60101689"  
/translation="MRGSELFTGVVPILVELDGDVNGHKFSVSGEGDATTYKLT  
KFICTTGKLPVPMPTLVTFYGYVQCARYPDHMKQHDFFKSAMPEGYVQERTIFEKD  
GNKYTRAEVKEEGDILVNRIELKIDFKEEDGNILGHLEYNYNSHNVTYADKQKNG  
IKVNFKIRHNIEDGSVQLADHYQONTPIGDVLPDNYHLYSTQALSKDPNEKRDHM  
VLEFVTAAGITTHGMDLYKSGSGCKTTASSAADIIVDGHSCFLGSGSGMSASE  
DVIKEFMRFKVRMEGSVNGHEFIEEGEGRPYEGTQAKLKVTKGGLPFAWDILSP  
QEQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDGAVTVTQDSSLQDGCPIYK  
VKFIGVNPSPDGPVMQKTMGWPESTERLYPRDGLVAKGEIHKALKLKDGGHYLVEFKS  
IYMAKKPVQLPGYYVVDKSLDITSHNEDYTIIVEQYRTERGRHLLFL"  
140\_.853  
/gene="GRFP"  
/note="cycle 3 mutant of GFP; derived initially from  
pGFPmut3.1; Region: green fluorescent protein mutant 3.1"  
854\_.952  
/gene="GRFP"  
/note="contains G,S residues for flexibility on each side,  
M13 forward and reverse primer sites, and NotI, EcoRV,  
Sali cloning sites; Region: linker"  
953\_.1627  
/gene="GRFP"  
/note="faster folding mutant of discosoma red fluorescent  
protein; derived from pDsRed-T3; silent mutations have  
been introduced to change restriction sites; Region:  
discosoma red fluorescent protein T3 mutant"  
1778\_.2638  
/function="ampicillin resistance"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="AA13968.1"  
/db\_xref="GI:60101688"  
/translation="MSIQHFRVALIPFPAFCLPVFAHPETLVKVKDAEDQLGARVGY  
IELDLSGKILESFRPEERFPMWSTFKVLGAVLSRIDAGEQQLGRRIRHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLITIGPKELTAFLLHNMGDHVTRL  
DRWEPELNEAIPNDERDITMPVAMATYLRKLLTGEELLTLASROQLIDMWEADKVAGPL  
LRNALPGWFIADKSGAGERSGRIIAALGPDGKPSRIVITYTGSQATWDERNRQIA  
EIGASLIKHW"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.16e-109 length: 3441  
Score: 1170.00 Matches: 217  
Percent Similarity: 97.78% Conservative: 3  
Best Local Similarity: 96.44% Mismatches: 5  
Query Match: 96.38% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY916793 (1-3441)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 953 ATGGCTCTCTCGAGGAGCTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGGAGGC 1012  
QY 21 ThrValaAngLYHISgluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 1013 TCCGTGAACGGCCACGAGTTTCAGATTCGAGGGCGAGGGCGGCCCCCTACGAGGCG 1072  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 1073 ACCGAGACCGCCAAGCTGAAGGTACCAAGGGCGGCCCTCGCCTTGGCTGGACATC 1132

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHISProAlaAspIlePro 80  
DB 1133 CTGTCCCCCCAGTTCACGTACGAGCTCCAAAGGTGTACGTGAAGCACCCCGCACATCCCC 1192  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLyTPrGluArgValMetAsnPhcGlu 100  
DB 1193 GACTACAAGAAGCTGTCTCTCCCGAGGCTTCAAGTGGAGCGCGTGATGAATTCAGAG 1252  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 1253 GACGCGCGCGTGTGACCGGTGACCCAGACTCTCTCCGACGAGACGGCTGCTCATCTAC 1312  
QY 121 LysValLySPheIleGlyValaAsnPheproSerAspGlyProValMetGlnLySLeuThr 140  
DB 1313 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAAATGCAAGAAGACT 1372  
QY 141 MetGlyTPrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 1373 ATGGGCTGGAGCCCTCCACCGAGCGCCTGTACCCCGCAGCGCGCTGTGAAGGGCGAG 1432  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 1433 ATCCACAAGGCCCTGAAGCTGAAGAGCGCGGCCACTACCTGTGAGTTCAGTTCATC 1492  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValaAspSerLySLeuAsp 200  
DB 1493 TACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACTAGTGACTCCAAGCTGGAC 1552  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 1553 ATCACTCCACACAAGAGACTACACCATCTGTGAGCAGTAGCAGCGCACCGAGGACGC 1612  
QY 221 HisHisLeuPheLeu 225  
DB 1613 CACCATCTGTCTCTT 1627  
RESULT 56  
AJ851285 5311 bp DNA circular SYN 12-NOV-2004  
LOCUS  
DEFINITION Cloning vector pRU1105 dsRedT3 gene for Ds Red fluorescent protein T3.  
ACCESSION AJ851285  
VERSION AJ851285.1 GI:55724876  
KEYWORDS Ds Red fluorescent protein T3; dsRedT3 gene.  
SOURCE Cloning vector pRU1105  
ORGANISM other sequences; artificial sequences; vectors.  
REFERENCE 1  
AUTHORS Karunakaran,R. and Poole,P.S.  
TITLE High throughput promoter probe vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 5311)  
AUTHORS Poole,P.S.  
TITLE Direct Submission  
JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of Reading, Whiteknights, Reading, RG6 6AJ, UNITED KINGDOM  
FEATURES  
source  
1. 5311  
/organism="Cloning vector pRU1105"  
/mol\_type="other DNA"  
/db\_xref="taxon:299180"  
/lab\_host="Escherichia coli"  
326\_.1003  
/gene="dsRedT3"  
326\_.1003  
/gene="dsRedT3"  
326\_.1003  
/gene="dsRedT3"  
/codon\_start=1  
/transl\_table=11  
/product="Ds Red fluorescent protein T3"  
/protein\_id="CAH64890.1"  
/db\_xref="GI:55724877"  
/translation="MASSEDVIKEFMRFKVRMEGSVNGHEFIEEGEGRPYEGTQTA  
KLKVTKGGLPFAWDILSPQEQYGSKVYVKHPADIPDYKKLSFPEGFKWERVMNFDG



GVTVTQDSSLQDGCFTYKVKFIGVNPSPDGPVMOKKTWGPSTERLYPRDGLKGE  
IHKALKDGGHYLVEFKSIYMAKKPVQLPGYVVDSKLDITSHNEDYTIIVEQYERTE  
GRHHLFL"

ORIGIN

Alignment Scores: 5,36e-109 Length: 5311  
Pred. No.: 1170.00 Matches: 217  
Score: 97.78% Conservative: 3  
Percent Similarity: 96.44% Mismatches: 5  
Best Local Similarity: 96.38% Indels: 0  
Query Match: 11 Gaps: 0  
DB:

US-10-006-922A-12 (1-225) x AJ851285 (1-5311)

QY 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLyValArgMetGluGly 20  
Db 326 ATGGCCTCCTCCGAGGACGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 385  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40  
Db 386 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC 445  
QY 41 HisAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60  
Db 446 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGCTGGACATC 505  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
Db 506 CTGTCCCCCAGTTCAGAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGCAGCATCCCC 565  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 566 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTGAG 625  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 626 GACGGCGCGGTGTGACCGTGACCCAGGACTCTCCCTGCAGGACGGCTGTCTCATCTAC 685  
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 686 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCCGACGGCCCCGTAATGCAGAAAGACT 745  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySgLyGlu 160  
Db 746 ATGGGCTGGAGCCCTCCACCGAGCGCGCTGTACCCCCCGACGGCGTGTGAAGGGCGAG 805  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
Db 806 ATCCACAAGGCCCTGAAGCTGAAGAGCGCGGCCACTACCTGTGAGTTCAATCCATC 865  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 866 TACATGGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGACTCCAAGCTGAGC 925  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 926 ATCACTCCCAACGAGGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 985  
QY 221 HisHisLeuPheLeu 225  
Db 986 CACCACCTGTTCCTG 1000

RESULT 57  
AJ851289 13079 bp DNA circular SYN 12-NOV-2004  
LOCUS Cloning vector PRU1075 gusA gene for beta-glucuronidase and dsRedT3  
DEFINITION gene for Ds Red fluorescent protein T3.  
ACCESSION AJ851289  
VERSION AJ851289.1 GI:55724885  
KEYWORDS beta-glucuronidase; Ds Red fluorescent protein T3; dsRedT3 gene;  
gusA gene.  
SOURCE Cloning vector PRU1075

ORGANISM Cloning vector PRU1075

other sequences; artificial sequences; vectors.

REFERENCE 1 Karunakaran,R. and Poole,P.S.  
AUTHORS High throughput promoter probe vectors  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 13079)  
REFERENCE Poole,P.S.  
AUTHORS Direct Submission  
TITLE Submitted (21-OCT-2004) Poole P.S., School of AMS, University of  
JOURNAL Reading, whiteknights, Reading, RG6 6AU, UNITED KINGDOM  
FEATURES location/Qualifiers  
source 1..13079  
1..13079  
/organism="Cloning vector PRU1075"  
/mol\_type="other DNA"  
/db\_xref="taxon:299184"  
/lab\_host="Escherichia coli"  
186..863  
/gene="dsRedT3"  
186..863  
/gene="dsRedT3"  
/gene="dsRedT3"  
/codon\_start=1  
/transl\_table=1  
/product="Ds Red fluorescent protein T3"  
/protein\_id="CAH64917.1"  
/db\_xref="GI:55724886"  
/translation="MASSEDIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEQTOTA  
KLKVTKGSPLEFAMDILSPQFOYGSKVYKHPADIPTYKLSPEEGFKWERVWNFEDG  
GVTVTQDSSLQDGCFTYKVKFIGVNPSPDGPVMOKKTWGPSTERLYPRDGLKGE  
IHKALKDGGHYLVEFKSIYMAKKPVQLPGYVVDSKLDITSHNEDYTIIVEQYERTE  
GRHHLFL"

gene

CDS

885..2696  
/gene="gusA"  
885..2696  
/gene="gusA"  
/codon\_start=1  
/transl\_table=1  
/product="beta-glucuronidase"  
/protein\_id="CAH64918.1"  
/db\_xref="GI:55724887"  
/translation="MVRPVEPTREIKKLDELGMAFSLDRENGCIDQRMMESALQESRA  
IAYGSEFNDOPADADIRNVAQVWVYQREVFIPKGWAGORIVLRFDAVTHYGKVVNNQ  
EVMHQGGYTPPEADVTPYIYAGKSVRITVCVNNELNMQITPBGAVITDENGKKQSY  
FHDFENYAGIHSVMLYTPPTWYDDITVTHVAODCNHASVDWQVANGDVSVELRD  
ADQVAVATGQSTGTLQVNVPHLMQPGEGYLVELCVTAKSOTBCDIYPLRVGIRSVAV  
KGEQFLINHKEPYFTGFGHEHDADLRGKGFDNVLAMVHDLAMDIGANSYRTSHYPYA  
EEMLDWADEHGIVIDETAAGFNLGLIGFEAGNKPKELYSBAVNGETQQAHLQAI  
KELIARDKNHPSVVMMSIANEPDTRPGAREYFAPLAETRKLDPTRPITCVNMFCD  
AHTDRTSDFDVLCLNRYGWWYQSDLETAEKVLLEKELIWAQEKLIHQPIITTEYGVN  
TLAGLSMYTDMWSESEYQCAWLDMTHRVFDRVSAVVEGQVWNPADFATSQGILRVGNN  
KKGIETRRDRKPKSAAFLLQKRWTGMNFGKPKQGGKQ"

ORIGIN

Alignment Scores:

Pred. No.: 1,61e-108 Length: 13079  
Score: 1170.00 Matches: 217  
Percent Similarity: 97.78% Conservative: 3  
Best Local Similarity: 96.44% Mismatches: 5  
Query Match: 96.38% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AJ851289 (1-13079)

QY 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLyValArgMetGluGly 20  
Db 186 ATGGCCTCCTCCGAGGACGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 245  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40  
Db 246 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC 305  
QY 41 HisAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60



Db 306 ACCCAGACCGCCAGCTGAAGGTGACCAAGGGCGGCCCTTCGCTTCGCGTGGAGACATC 365  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIlySHsPProAlaAsp1IePro 80  
366 CTGTCCCCCAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 425  
QY 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGlu 100  
426 GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 485  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe1IeTyr 120  
486 GACGGCGCGGTGTGACCGGTGACCCAGAACTCTCTCCCTGACGACGGCTGCTCATCTAC 545  
QY 121 LysValLysPhe1IeGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
546 AAGGTGAAGTTTCATCGCGCTGAATTTCCCTCCGACGGCGCCCGTAATGCAGAGAAGACT 605  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
606 ATGGGCTGGAGACCCCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGGCGAG 665  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSer1Ie 180  
666 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGTGAAGTTCAGTCCATC 725  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
726 TACATGGCCACAAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGACTCCAAGCTGAC 785  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
786 ATCACTCTCCACAACGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCG 845  
QY 221 HisHisLeuPheLeu 225  
Db 846 CACCACCTGTTCTTG 860

RESULT 58  
LOCUS AJ851286 5311 bp DNA circular SYN 12-NOV-2004  
DEFINITION Cloning vector pRU1106 dsRedT4 gene for Ds Red fluorescent protein T4.  
ACCESSION AJ851286  
VERSION AJ851286.1 GI:55724878  
KEYWORDS Ds Red fluorescent protein T4; dsRedT4 gene.  
SOURCE Cloning vector pRU1106  
ORGANISM other sequences; artificial sequences; vectors.  
REFERENCE 1  
AUTHORS Karunakaran,R. and Poole,P.S.  
TITLE High throughput promoter probe vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 5311)  
AUTHORS Poole,P.S.  
TITLE Direct Submission  
JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of Reading, Whiteknights, Reading, RG6 6AU, UNITED KINGDOM  
FEATURES  
Source location/Qualifiers  
1..5311  
/organism="Cloning vector pRU1106"  
/mol\_type="other DNA"  
/db\_xref="taxon:299181"  
/lab\_host="Escherichia coli"  
326..1003  
/gene="dsRedT4"  
326..1003  
/gene="dsRedT4"  
326..1003  
/gene="dsRedT4"  
/codon\_start=1  
/transl\_table=11  
/product="Ds Red fluorescent protein T4"  
/protein\_id="CAH64891.1"  
/db\_xref="GI:55724879"

/translation="MASSEDAVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEQTQTAKLKVTKGGPLPFAWDILSPQFYGSKYVYKHADI PDYKKLSFPEGFKMERVMNFEDG GVVTVTQDSLDGCFIYKVKFI GVNFPSDGFVMOKTMGMEPSTERLYPRDGYLKE IHKALKKDGGHYLVBFKSIYMAKKPVQLPGYYVDSKLDITSHNEDYTIIVEQYERAE GRHHLFL"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.74e-108 Length: 5311  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatch: 6  
Query Match: 95.96% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AJ851286 (1-5311)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 326 ATGGCCTCTCCGAGAGCGTCATCAAGAGCTTCATGCGCTTCAAGGTGCGCATGAGGCG 385  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 386 TCCGTGAACGCCACGAGTTCGAGATCGAGGCGGAGGCGGAGGCGCCCTTACGAGGCG 445  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAsp1Ie 60  
Db 446 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGCGGCCCTTCCTTCGCTTGGAGACATC 505  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisPProAlaAsp1IePro 80  
Db 506 CTGTCCCCCAGTTCAGTACGAGCTCCAAAGGTGTACGTGAAGCACCCCGCGACATCCCC 565  
QY 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 566 GACTACAAGAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 625  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPhe1IeTyr 120  
Db 626 GACGGCGCGGTGTGACCGGTGACCCAGACTCTCTCCGACGAGCGGCTTCAATCTAC 685  
QY 121 LysValLysPhe1IeGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 686 AAGGTGAAGTTTCATCGCGCTGAATTTCCCTCCGACGCGCCGTAATGCAGAGAAGACT 745  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 746 ATGGGCTGGAGCCCTCCACCGAGCGCTGTACCCCGCGACGCGGTGTGAAGGGCGAG 805  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSer1Ie 180  
Db 806 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC 865  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 866 TACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAGCTGAGC 925  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 926 ATCACTCTCCACAACGAGACTACCATCTGTGAGCAGTACGAGCGCGCGAGGGCGCG 985  
QY 221 HisHisLeuPheLeu 225  
Db 986 CACCACCTGTTCTTG 1000

RESULT 59  
LOCUS AY342347 10141 bp DNA linear SYN 21-OCT-2003  
DEFINITION Red H-Pelican DsRed.T4 transformation vector, complete sequence.  
ACCESSION AY342347  
VERSION AY342347.1 GI:33358309  
KEYWORDS Red H-Pelican DsRed.T4 transformation vector  
SOURCE



ORGANISM Red H-Pelican DsRed.T4 transformation vector  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 10141)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP [DsRed.T4] and Nuclear RFP in Insulated Vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 10141)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-2003) Division of Biology/CDB, UCSD, 4121 Bonner Hall, MC 0349, La Jolla, CA 92093-0349, USA  
FEATURES  
source 1..10141  
/organism="Red H-Pelican DsRed.T4 transformation vector"  
/mol\_type="other DNA"  
/db\_xref="taxon:240544"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,82e-108 Length: 10141  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.96% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY342347 (1-10141)  
QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 285 ATGGCCTCCTCCGAGACGTCATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGGC 344  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 345 TCCGTGAACGGCCAGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 404  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 405 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGCGCCCTGCGCTGGGACATC 464  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 465 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTAAGTGAAGCACCCTGACATCCCC 524  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 525 GACTACAAAGAGCTGCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTTCGAG 584  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 585 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGACAGACGGCTGCATCTAC 644  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 645 AAGGTGAAGTTCAATCGCGGTGAATCTCCCTCCGACGGCCCCGTAAATGCAAGAAGACT 704  
QY 141 MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 705 ATGGGCTGGAGCCCTTCCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 764  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySeriIle 180  
Db 765 ATCCACAAAGCCCTGAAGCTGAAGAGACGGCGGCCACTAATCTGTGAGTTCAAGTTCATC 824  
QY 181 TyrMetAlaLySLeSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 825 TACATGGCCAAGAGCCCGTGCAGCTGCCCGCTACTACTAAGTCCAAAGCTGGAC 884  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 885 ATCACTCCCAACGAGAGCTACACCATCTGTGAGAGTACGAGCGCGCCGAGGGCGGC 944

QY 221 HisHisLeuPheLeu 225  
Db 945 CACCACTGTCTCTG 959  
RESULT 60  
AY342348 10276 bp DNA linear SYN 21-OCT-2003  
LOCUS Red H-Stinger DsRed.T4-NLS transformation vector, complete  
DEFINITION  
ACCESSION AY342348  
VERSION AY342348  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 10276)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP [DsRed.T4] and Nuclear RFP in Insulated Vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 10276)  
AUTHORS Barolo,S. and Posakony,J.W.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-2003) Division of Biology/CDB, UCSD, 4121 Bonner Hall, MC 0349, La Jolla, CA 92093-0349, USA  
FEATURES  
source 1..10276  
/organism="Red H-Stinger DsRed.T4-NLS transformation vector"  
/mol\_type="other DNA"  
/db\_xref="taxon:240545"  
ORIGIN

Alignment Scores:  
Pred. No.: 3,88e-108 Length: 10276  
Score: 1165.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Best Local Similarity: 96.00% Mismatches: 6  
Query Match: 95.96% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AY342348 (1-10276)  
QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 288 ATGGCCTCCTCCGAGACGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 347  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 348 TCCGTGAACGGCCAGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 407  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 408 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGCGCCCTGCGCTGGGACATC 467  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 468 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGTAAGTGAAGCACCCTGACATCCCC 527  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 528 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 587  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 588 GACGGCGCGGTGTGACCGTGAACCAAGACTCTCTCCCTGACAGACGGCTGCATCTAC 647  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 648 AAGGTGAAGTTCAATCGCGGTGAATCTCTCCCGACGGCCCCGTAAATGCAAGAAGACT 707  
QY 141 MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160



Db 708 ATGGGCTGGAGACCTCCACCGAGCGCCTGTACCCCGCGCAGCGCTGTGAAGGCGAG 767

Qy 161 ILeHISLysAlaLeuLysLeuLysAspGlyGlyHISLysLeuValGluPheLysSerIle 180

Db 768 ATCCACAAGGCCCTGAAGCTGAAGACGGCGCCACTACTACTGAGTGAAGTTCAGATCCATC 827

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 828 TACATGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGAAGTCCAGAGCTGAC 887

Qy 201 IleThrSerHISAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 888 ATCACTCTCCCAACAAGAGGACTACACCATCTGTGAGCAGTACGAGCGCGCCGAGGCGCC 947

Qy 221 HISHISLeuPheLeu 225

Db 948 CACCACCTGTTCCTG 962

RESULT 61

LOCUS AY490568 10481 bp DNA linear SYN 24-DEC-2003

DEFINITION UAS-Red Stinger DsRed.T4-NLS Drosophila UAS-RF transformation vector, complete sequence.

ACCESSION AY490568

VERSION AY490568.1 GI:40218087

KEYWORDS UAS-Red Stinger DsRed.T4-NLS Drosophila UAS-RF transformation vector

SOURCE UAS-Red Stinger DsRed.T4-NLS Drosophila UAS-RF transformation vector

ORGANISM UAS-Red Stinger DsRed.T4-NLS Drosophila UAS-RF transformation vector

REFERENCE 1 (bases 1 to 10481)

AUTHORS Barolo,S. and Posakony,J.W.

TITLE Advances in Drosophila Transgenic Reporters I: Fast-Maturing RFP [DsRed.T4] and Nuclear RFP in Insulated Vectors

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 10481)

AUTHORS Barolo,S. and Posakony,J.W.

TITLE Direct Submission

JOURNAL Submitted (13-NOV-2003) Cell & Developmental Biology, University of Michigan Medical School, 5732 Med Sci 2, Ann Arbor, MI 48109, USA

FEATURES

source

1. 10481

/organism="UAS-Red Stinger DsRed.T4-NLS Drosophila UAS-RF transformation vector"

/mol\_type="other DNA"

/db\_xref="taxon:258595"

ORIGIN

Alignment Scores:

pred. No.: 3.98e-108 length: 10481

Score: 1165.00 Matches: 216

Percent Similarity: 97.33% Conservative: 3

Best Local Similarity: 96.00% Mismatches: 6

Query Match: 95.96% Indels: 0

DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AY490568 (1-10481)

Qy 1 MecArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20

Db 494 ATGGCTCTCTCCGAGGACGTCAAGAGTTCATGCGCTCAAGGTGCGCATGAGGCG 553

Qy 21 ThrValAsnGlyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 554 TCCGTGAACGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGCG 613

Qy 41 HIsAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60

Db 614 ACCCAGACCCGCCAAGCTGAAGGTGACCAAGGGCGGCGCCCTCGCCTGGGACATC 673

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHISProAlaAspIlePro 80

Db 674 CTGTCCCCCAGTTCAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 733

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100

Db 734 GACTACAAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 793

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 794 GACGGCGCGTGGTGAACCGTGAACCCAGACTCTCTCCCTGCAGAGCGGCTGTCAATCAC 853

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140

Db 854 AAGGTGAAGTTCATCGCGGTGAAGTTCCTCCGACGGCCCCGTATGAGAAGAGACT 913

Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 914 ATGGGCTGGAGCCCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 973

Qy 161 ILeHISLysAlaLeuLysLeuLysAspGlyGlyHISLysLeuValGluPheLysSerIle 180

Db 974 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGAGTTCAGTTCATC 1033

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

Db 1034 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGTCCAAAGTTCGAC 1093

Qy 201 IleThrSerHISAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 1094 ATCACTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCGCGAGGCGCC 1153

Qy 221 HISHISLeuPheLeu 225

Db 1154 CACCACCTGTTCCTG 1168

RESULT 62

LOCUS AJ851290 13079 bp DNA circular SYN 12-NOV-2004

DEFINITION Cloning vector pRU1076 gusA gene for beta-glucuronidase and dsRedT4 gene for Ds Red fluorescent protein T4.

ACCESSION AJ851290

VERSION AJ851290.1 GI:55724888

KEYWORDS beta-glucuronidase; Ds Red fluorescent protein T4; dsRedT4 gene; gusA gene.

SOURCE Cloning vector pRU1076

ORGANISM Cloning vector pRU1076

REFERENCE 1

AUTHORS Karunakaran,R. and Poole,P.S.

TITLE High throughput promoter probe vectors

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 13079)

AUTHORS Poole,P.S.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of Reading, Whiteknights, Reading, RG6 6AJ, UNITED KINGDOM

FEATURES

source

1. 13079

/organism="Cloning vector pRU1076"

/mol\_type="other DNA"

/db\_xref="taxon:299185"

/lab\_host="Escherichia coli"

186. .863

/gene="dsRedT4"

186. .863

/gene="dsRedT4"

/gene="dsRedT4"

/codon\_start=1

/transl\_table=11

/product="Ds Red fluorescent protein T4"

/protein\_id="CAH64919.1"

/db\_xref="GI:55724889"

/translation="MASSEDVIKEFMRFKVRMEGSVNGHEFIEGEGGRPYEGTQTA KLKVTKGGRPLFAMDLISPOFGSKYVYKHADIPDYKGLSPFEGFKRVRVMNFEDG GVTVTQDSLQDGCPIYKVKFIGVNPSPDGPVMQKTMGWPSTERTLYPRDVLKGE



IHKALKLKDGHYLVEFKSIYMAKPVQLPGYYVDSKLDITSHNEDYTIVEQYERAE
 GRHHLFL"
 gene
 885..2696
 /gene="gusa"
 CDS
 885..2696
 /gene="gusa"
 /codon\_start=1
 /transl\_table=11
 /product="beta-glucuronidase"
 /protein\_id="CAH64920.1"
 /db\_xref="GI:55724890"
 /translation="MVRPVEPTREIKLDELWAFSLDRENCGIDQRMWESALQESRA
 IAVPGSFNDQFADADIRNAGNVWYQREVFI PKMAGRIVLRFDAVTHYKQVWVNNQ
 EVMHQGGYTPFEADVTPYIAGKSVRTVCVNNELNWQIIPGMVITDENGKKQSY
 FHDFFNYAGIHRVMLYTPNTWVDITVTVHVAQDCNHASVDWQVANGDVSELRD
 ADQOVVATGGTSGTLQVNPMLMQEGYLVELCTVAKSQTECDIYPLRVGISVAV
 KGEQFLINHKEPFYFTGFGRHEDADLRKGFNDVLMVHDHALMDWIGANSYRTSHPYA
 EEMLDWADEHGIIVIDETA VFNLSLIGFEAGNKPKELYSEAVNGETQOAHQAI
 KELIARDKNHPSVVMWSIANEPDTRPQAREYFAPLAEATRKLDPTRPITCVNVMECD
 AHTDTISDLFDVLCINRYGVWVQSGDLEAKVLEKELAWQEKLHQPIITTEGYD
 TLAGLSMYTDMWSEBYOCAMDYHRVFDRSVAVGEQWNFADFATSQGILRVGN
 KKGIFTRDRKPKSAAPFLQKRWGMNGEKPQGGKQ"

Alignment Scores:
 Pred. No.: 5.21e-108 Length: 13079
 Score: 1165.00 Matches: 216
 Percent Similarity: 97.33% Conservative: 3
 Best Local Similarity: 96.00% Mismatches: 6
 Query Match: 95.96% Indels: 0
 DB: 11 Gaps: 0

ORIGIN

US-10-006-922A-12 (1-225) x AJ851290 (1-13079)
 QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
 DB 186 ATGGCCTCCTCCGAGGACGTGCATCAAGGAGTTTCATCGCTTCAAGGTGCGCATGAGGGC 245
 QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
 DB 246 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCCCTTACGAGGGC 305
 QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
 DB 306 ACCCAGACCCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTTCCGCTGGGACATC 365
 QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
 DB 366 CTGTCCCCCAGTTCAGTA CGGCTCCAAGGTGTA CGTGAAGCA CCCCCCAGCATCCCC 425
 QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
 DB 426 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAATTCGAG 485
 QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
 DB 486 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCGACGAGCGGCTGTCACTCTAC 545
 QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
 DB 546 AAGGTGAAGTTTCATCGGCTGAATCTTCCCTCCGACGGCCCCCGTATATGCAAGAGAAGACT 605
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
 DB 606 ATGGGCTGGGAGCCCTCCACCGAGCGCTGTACCCCCCGACGGCGTGTGAAGGGCGAG 665
 QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
 DB 666 ATCCACAAGGCCCTGAAGCTGAAGACGCGGCGGCACACTACTGTGTGAGTTCAAGTTCATC 725
 QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
 DB 726 TACATGGCCAAGAAGCCCGTGAAGCTGCGCGGCTACTACTACTGACTCCAAAGCTGAC 785

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
 DB 786 ATCACTCCACCAACGAGACTACCATCTGTGAGCAGTACGAGCGCCGAGGGCCG 845
 QY 221 HisHisLeuPheLeu 225
 DB 846 CACCACCTGTTCTCTG 860

RESULT 63
 AF506025 681 bp DNA linear SYN 19-JUN-2002
 LOCUS AF506025
 DEFINITION Synthetic construct dimeric red fluorescent protein gene, complete
 ACCESSION AF506025
 VERSION AF506025.1 GI:21464833
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 12060735
 2 (bases 1 to 681)
 Campbell,R.E., Tour,O., Palmer,A.E., Steinbach,P.A., Baird,G.S.,
 Zacharias,D.A. and Tsien,R.Y.
 A monomeric red fluorescent protein
 Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7877-7882 (2002)

FEATURES
 source
 1..681
 /organism="synthetic construct"
 /mol\_type="genomic DNA"
 /db\_xref="taxon:32630"
 /note="from Discosoma sp.; Clontech vector pDsRed-N1 with
 human optimized codon usage"
 1..681
 /note="dimer2; engineered variant of tetrameric red
 fluorescent protein from Discosoma deposited in GenBank
 Accession Number AF03369"
 /codon\_start=1
 /transl\_table=11
 /product="dimeric red fluorescent protein"
 /protein\_id="AAM54542.1"
 /db\_xref="GI:21464834"
 /translation="MVASSEDVYKEFMRFKVRMEGSVNGHEFIEGEGGRPYEGTQT
 AKLKYTKGGLPFAWDILSPQYGSKAYVKHPADIPYKLSFPEGFKWERVMNFED
 GGVTVTQDSSLQDGTLIYKVKFRGTNPPDPGPMQKTMGEASTERLYPRDGLVKG
 EIHQALKLKDGHYLVEFKTIYMAKPVQLPGYYVDTKLDITSHNEDYTIVEQYERS
 EGRHHLFL"

CDS
 1..681
 /note="dimer2; engineered variant of tetrameric red
 fluorescent protein from Discosoma deposited in GenBank
 Accession Number AF03369"
 /codon\_start=1
 /transl\_table=11
 /product="dimeric red fluorescent protein"
 /protein\_id="AAM54542.1"
 /db\_xref="GI:21464834"
 /translation="MVASSEDVYKEFMRFKVRMEGSVNGHEFIEGEGGRPYEGTQT
 AKLKYTKGGLPFAWDILSPQYGSKAYVKHPADIPYKLSFPEGFKWERVMNFED
 GGVTVTQDSSLQDGTLIYKVKFRGTNPPDPGPMQKTMGEASTERLYPRDGLVKG
 EIHQALKLKDGHYLVEFKTIYMAKPVQLPGYYVDTKLDITSHNEDYTIVEQYERS
 EGRHHLFL"

ORIGIN
 Alignment Scores:
 Pred. No.: 4.42e-105 Length: 681
 Score: 1121.00 Matches: 207
 Percent Similarity: 95.96% Conservative: 7
 Best Local Similarity: 92.83% Mismatches: 9
 Query Match: 92.34% Indels: 0
 DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x AF506025 (1-681)
 QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22
 DB 10 TCCTCCGAGGACGTATCAAAAGATTTCATGCGCTTCAAGGTGCGCATGAGGCTCCGTG 69
 QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsn 42
 DB 70 AACGGCCACGAGTTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGCAACCCAG 129
 QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62



Db	130	ACCGCCAAGCTGAAGGTGACCAGAAGGGCGGCCCCCTGCCTTCGGCCTGGGACATCTGTC	189
QY	63	ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr	82
Db	190	CCCCAGTTCCAGTAGCGGCTCCAAAGGCGTGACTGAAGCACTCCGCCGACTAC	249
QY	83	LysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGluAspGly	102
Db	250	AAGAAGCTGTCTTCCCAGAGGGCTTCAAGTGGAGCGGTGATGAATTGAGGACGCG	309
QY	103	GlyValValThrValThrglnAaspSerSerLeuGlnAaspGlyCysPheIleTyrLysVal	122
Db	310	GGCGTGATGACCGTGACCAGACTCCTCCCTGCAGACGGCACGCTGATCTACAAGTG	369
QY	123	LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly	142
Db	370	AAGTTCGCGGACCACTTCCCCCGACGGGCCCGTAATGCAAGAAGAACCATGGGC	429
QY	143	TyrGluAlaSerThrgluArgLeuTyrProArgAspGlyValLeuLysGlyLileHis	162
Db	430	TGGAGGCTCCACCGAGCGCTGTACCCCCCGCAGCGGCTGTGAAGGGCGAGATCAC	489
QY	163	LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet	182
Db	490	CAGGCCCTGAAGCTGAAGAGACGGCGGCACCTACTGTGTGAAGTTCAAGACCATCTAC	549
QY	183	AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr	202
Db	550	GCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACACCAAGCTGGACATCAC	609
QY	203	SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis	222
Db	610	TCCACAAACGAGACTACACCATCTGTGAACAGTACGAGCGCTCCGAGGGCGGCACAC	669
QY	223	LeuPheLeu 225	
Db	670	CTGTTCTTG 678	
RESULT 64			
LOCUS	AF506026	1395 bp	DNA linear SYN 19-JUN-2002
DEFINITION	Synthetic construct tandem-dimer red fluorescent protein gene, complete cds.		
ACCESSION	AF506026		
VERSION	AF506026.1	GI:21464835	
KEYWORDS	synthetic construct synthetic construct other sequences; artificial sequences.		
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 (bases 1 to 1395) Campbell,R.E., Tour,O., Palmer,A.E., Steinbach,P.A., Baird,G.S., Zacharias,D.A. and Tsien,R.Y. A monomeric red fluorescent protein Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7877-7882 (2002)		
TITLE	JOURNAL PUBLISHED		
REFERENCE	2 (bases 1 to 1395) Campbell,R.E. and Tsien,R.Y. Direct Submission Submitted (24-APR-2002) Pharmacology, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA		
JOURNAL			
FEATURES			
source	Location/Qualifiers		
	1..1395		
	/organism="synthetic construct"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32630"		
	/note="from Discosoma sp.; Clontech vector pDsRed-N1 with human optimized codon usage"		
	1..1395		
	/note="cdimer2(12); engineered variant of tetrameric red fluorescent protein from Discosoma deposited in GenBank Accession Number AAF03369"		
	/codon_start=1		
CDS			

	/transl_table=11
	/product="tandem-dimer red fluorescent protein"
	/protein_id="AAM54543.1"
	/db_xref="GI:21464836"
	/translation="MVASSEDVIKEFMRPKVRMEGVNGHEFEIEGEGRPYEQTOTAKLVTKGGPLPAMDILSPQYGSKAYVKHPADIPDYKLSPPEGFKMERVMNFDGGVTVTODSLDGTLYKKVFRGTNPDPGVMOCKTMGEASTERLYPRDGLKGFIHQALKLKDGHYLVEFTIYMAKKPVLPGYYVDITKLDTSHNEDYTIIVQYERSBGRHHLFL"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.06e-104 Length: 1395
Score:	1121.00 Matches: 207
Percent Similarity:	95.96% Conservative: 7
Best Local Similarity:	92.83% Mismatches: 9
Query Match:	92.34% Indels: 0
DB:	Gaps: 0
US-10-006-922A-12 (1-225) x AF506026 (1-1395)	
OY	3 SerSerLyAsnValIllelysgluPhmetArgphelysValArgmetgluIyThrVal 22      :::
Db	10 TCCTCGAGGACGTCAAGAAGTCATGCCTTCAAGTGCGCATGAGGGCTCCGTG 69    :::
OY	23 AsnglYHisgluPhesgluilegluylgluIygluIyArgProtyrGlunghisAsn 42      :::
Db	70 AACGCCACGAGTTGAGATCGAGGCCAGGGCGAGGGCCCCCTACGAGGCCACCAG 129    :::
OY	43 ThrVallysleuLysValThrlYsglyglyProleuprophelaTrpasrilleuSer 62      :::
Db	130 ACCGCCAAGCTGAAGGTGACCAAGGGGGCCCCCTGCCCTTGGAATCCTGTCC 189    :::
OY	63 ProclnpheglntyrGlyserLyasValTyrrValLysHisproalaasprieproaspTy 82      :::
Db	190 CCCCAgTTCAGTAGCGGCTCCAAGCGGTACGTGAAGCACCCCGCGACATCCCAGACTAC 249    :::
OY	83 LyeLysleuSerPheProgluGlyPheLysTrpgluArgvalMetAsnPhegluAspGly 102      :::
Db	250 AAGAAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAGAGACGGC 309    :::
OY	103 GlYValValThrValThrglnAspSerSerLeuglnAspGlyCySpheIIeTYrLysVal 122      :::
Db	310 GGCgtGTGACCGTGACCCAGACTTCTCTGACAGACGGCACGCTGATCTACAAGGTG 369    :::
OY	123 LySpheIIeglyValAsnpheProseerAspGlyProvalmetGlnlyLysThrmcGly 142      :::
Db	370 AAGTTCGCGGACCAACTTCCCCCGAGCGGCCCGTAATGCAGAAAGAACCATTGGGC 429    :::
OY	143 TrpgluAlaserThrgluArgleutyrrProargAspGlyValleuLysgluIleHis 162      :::
Db	430 TGGGAGGCTTCCACCGAGCGCTGTACCCCCGCGAGCGGTGTGAAGGGGAGATCCAC 489    :::
OY	163 LysAlaleuLysleuLysAspGlyglYhisTYrleuValgluPhelysSerIleTyrrMet 182 :::     :::
Db	490 CAGGCTTGAAGCTGAAGGACGGCGGCCACTACCTGTGTGAGTTCAAGACCATCTACATG 549    :::
OY	183 AlAlayLysProValGlnleuProglYTyrrTyrrValAspSerLySleuAsprieThr 202      :::
Db	550 GCCAAGAAGCCCGTGACGTGCCCGGCTACTACTACGTGACACCAAGCTGACATCAC 609    :::
OY	203 SerHisangluAspTyrrThrIleValgluGlnIntyrrgluArgThrgluIyARghisHis 222      :::
Db	610 TCCCAACAAGAGACTACACCATGTGGAACAGTAGAGCGCTCCGAGGGCGCCACACAC 669    :::
OY	223 Leupheleu 225      :::
Db	670 CTGTTCTTG 678    :::



```

RESULT 65
LOCUS      AX824729                      678 bp    DNA          linear      PAT 11-DEC-2003
DEFINITION Sequence 11 from Patent WO02068459.
ACCESSION  AX824729
VERSION     AX824729.1  GI:39750593
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
REFERENCE   1
AUTHORS     Non aggregating fluorescent proteins and methods for using the same
TITLE       Patent: WO 02068459-A 11 06-SEP-2002;
JOURNAL     Location/Qualifiers
FEATURES
    source
        1..678
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="hybrid coding sequence"
ORIGIN
Alignment Scores:
Pred. No.:      7.03e-105      Length:      678
Score:          1119.00        Matches:     206
Percent Similarity: 96.89%      Conservative: 12
Best Local Similarity: 91.56%    Mismatches:   7
Query Match:    92.17%         Indels:      0
DB:             Gaps:         0

US-10-006-922A-12 (1-225) x AX824729 (1-678)
QY      1 MetArgSerSerLybAsnValIleLybGluPheMetArgPheLybValArgMetGluLy 20
Db      1 ATGAGCTGCAGCAAGAACTGATCAAGGAGTTTCATCGGTTCAAGGTGCGGATGAGGGC 60
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugLy 40
Db      61 ACCGTGAACGGCCACGAGTTTCAGATCAAGGGCGAGGGCGAGGGCCGCCCTTACGAGGGC 120
QY      41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
Db      121 CACTGCAGCGTGAAGCTCATGTGACCAAGGGCGGCCCTCCCTCGCTTCGACATC 180
QY      61 LeuSerProGlnPheGlnTyrgLySerLybValTyrgValLybHisProAlaAspIlePro 80
Db      181 CTCAGCCCCCAGTTCCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 240
QY      81 AspTyrgLybLybLeuSerPheProGluGlyPheLybTyrgLyArgValMetAsnPheGlu 100
Db      241 GACTACAAAGAACTCACTTCCCGAGGGCTTCAAGTGGGAGCGGGTGATGAACCTTCGAG 300
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTy 120
Db      301 GACGGCGCGTGTGACCGTGAGCCAGACAGCAGCCTCAAGAGACGGCTGCTTCATCTAC 360
QY      121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140
Db      361 GAGGTGAAGTTCACTCGGCTGAACTTCCCGACGACGGCCCCGTGATGACGCGCGGACC 420
QY      141 MetGlyTyrgLybAlaSerThrGluArgLeuTyrgProArgAspGlyValLeuLybGlyGlu 160
Db      421 CGGGGCTGGAGGCGCCAGCAGCAGCGGCTTACCCCCGGAGCGGCTGCTCAAGGGCGGAC 480
QY      161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrgLeuValGluPheLybSerIle 180
Db      481 ATCCACATGGCCCTCCGGCTCGAGGGCGGCGGCCACTACCTCGTGAGTTCAAGACATC 540
QY      181 TyrgMetAlaLybLybProValGlnLeuProGlyTyrgTyrgValAspSerLybLeuAsp 200
Db      541 TACATGGCCAAAGAGCCCGTGCAGCTCCCGGCTACTACTACGTGAGACAGCAAGCTCGAC 600
QY      201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgLyArgThrGluGlyArg 220
```

```

Db      601 ATCACCAAGCCACAACGAGACTACCATCGTGCAGCAGTACGAGCGGACCGAGGGCCGG 660
QY      221 HisHisLeuPheLeu 225
Db      661 CACCACTCTTCCTC 675

RESULT 66
LOCUS      AY678268                      705 bp    DNA          linear      SYN 17-DEC-2004
DEFINITION Synthetic construct dimeric red fluorescent protein gene, complete cds.
ACCESSION  AY678268
VERSION     AY678268.1  GI:55420618
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
REFERENCE   1
AUTHORS     Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N.,
            Palmer,A.E. and Tsien,R.Y.
TITLE       Improved monomeric red, orange and yellow fluorescent proteins
JOURNAL     Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED     15558047
REFERENCE   2 (bases 1 to 705)
AUTHORS     Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and
            Tsien,R.Y.
TITLE       Direct Submission
JOURNAL     Submitted (06-JUL-2004) Pharmacology, University of California, San
            Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
FEATURES
    source
        1..705
            /organism="synthetic construct"
            /mol_type="other DNA"
            /db_xref="taxon:32630"
            /note="dtomato; engineered variant of dimeric red
            fluorescent protein dimer2 in GenBank Accession Number
            AF506025; N- and C-terminal replaced with equivalent
            residues from EGFP"
            /codon_start=1
            /transl_table=11
            /product="dimeric red fluorescent protein"
            /protein_id="AAV52168.1"
            /db_xref="GI:55420619"
            /translation="MVSKEGEVYKEFMRFKVRMEGSMNGHEFEIEGEGRPYEQTOT
            AKLKYTKGGPLPFAWDILSPQFMYSKAYVKHPADIPDYKLSFPEGFKMERVMNFED
            GGLVTVTQDSSLQDGLIYKVVMRGTFNPPDGPVMOKTIMGWASTERLYPRDGLKGG
            EIHQALKDKDGGHYLVBEFKTIYNAKKPVQLPGYYVYDTKLDITSHNEDYTIIVEQYERS
            EGRHHLFLYGMDELTK"
CDS
1..705
/note="dtomato; engineered variant of dimeric red
fluorescent protein dimer2 in GenBank Accession Number
AF506025; N- and C-terminal replaced with equivalent
residues from EGFP"
Alignment Scores:
Pred. No.:      2.08e-102      Length:      705
Score:          1095.00        Matches:     201
Percent Similarity: 94.57%      Conservative:  8
Best Local Similarity: 90.95%    Mismatches:   12
Query Match:    90.20%         Indels:      0
DB:             Gaps:         0

US-10-006-922A-12 (1-225) x AY678268 (1-705)
QY      5 LybAsnValIleLybGluPheMetArgPheLybValArgMetGluGlyThrValAsnGly 24
Db      16 GAGGAGTCAATCAAGAGTTTCATGCGCTTCAAGCTGCGCATGAGGGCTCCATGAACGGC 75
QY      25 HisGluPheGluIleGluGlyGluGlyArgProTyrgLybGlyHisAsnThrVal 44
Db      76 CACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCTGCCCTACGAGGGCACCAACGACCGCC 135
QY      45 LybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGln 64
```



Db	136	AAGCTGAAGGTGACCAAGGCGGCCCCCTTGCCCTTGCGCTGGACATCTGTCCCCCAG	195
Qy	65	PheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLys	84
Db	196	TTTCATGTAAGGCTCCAAAGCGGTACGTGAAGCAACCCCGCAGCATCCCCGATTACAAGAG	255
Qy	85	LeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyVal	104
Db	256	CTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAGAGACGGCGGTCTG	315
Qy	105	ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPhe	124
Db	316	GTGACCGTGAACCAAGACTCTCTCTGCAAGGACGCGCATCTTAACAAGGTGAAGATG	375
Qy	125	IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu	144
Db	376	CGCGGACCAACTTCCCCCGACGCGCCCGTAAATGCAAGAAGACCATGGCTGGAG	435
Qy	145	AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAla	164
Db	436	GCCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAGATCCACCAAGGCC	495
Qy	165	LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLys	184
Db	496	CTGAAGCTGAAGAGACGCGCGCCACATACCTGTGAGTTCAAGACCATCTACATGCGCAAG	555
Qy	185	LysProValGlnLeuProGlyTyrTyrTyrValAlaSerLysLeuAspIleThrSerHis	204
Db	556	AAAGCCGTGCAACTGCCCCGGCTACTACTACGTGACACCAAGCTGACATCACCTCCAC	615
Qy	205	AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHisLeuPhe	224
Db	616	AACGAGACTACACCATCGTGAACAGTACGAGCGCTCCGAGGCGCCACCACTGTTC	675
Qy	225	Leu 225	
Db	676	CTG 678	

RESULT 67			
LOCUS	AY678269	1431 bp	DNA linear SYN 17-DEC-2004
DEFINITION	Synthetic construct tandem-dimer red fluorescent protein gene, complete cds.		
ACCESSION	AY678269		
VERSION	AY678269.1	GI:55420622	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1 (bases 1 to 1431)		
AUTHORS	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N., Palmer,A.E. and Tsien,R.Y.		
TITLE	Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein		
JOURNAL	Nat. Biotechnol. 22 (12), 1567-1572 (2004)		
PUBMED	15558047		
REFERENCE	2 (bases 1 to 1431)		
AUTHORS	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and Tsien,R.Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUL-2004) Pharmacology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA		
FEATURES	Location/Qualifiers		
source	1..1431		
	/organism="synthetic construct"		
	/mol_type="other DNA"		
	/db_xref="taxon:32630"		
CDS	1..1431		
	/note="tdTomato; engineered variant of dimeric red fluorescent protein dimer2 in GenBank Accession Number AF506025; N- and C-termini replaced with equivalent residues from EGFP"		
	/codon_start=1		

/transl_table=1			
/product="tandem-dimer red fluorescent protein"			
/protein_id="AAV52169.1"			
/db_xref="GI:55420623"			
/translation="MYSKGEVKEFMRFKVRMEGSMNGHEFIEGEGRPYEGTOTAKLKVTKGGPLPPAWDILSPQFMYSKAYVKHPADI PDYKLSFPEGFKMERVMNFEDGGLVTVTQDSSLQDGTLLYKVKMRGTFNPPDGPVMQKTMGWEASTERLYPRDGYLKEBIHQALXLDKGHYLVEFKTIYMAKKPVQLPGYYVDTKLDITSHNEDYTIVEQYERS EGRHHLFLHGHTSGTSGSSSGTASSEDNNMAVIKEFMRFKVRMEGSMNGHEFIEGEGKMERVMNFEDGGLVTVTQDSSLQDGTLLYKVKMRGTFNPPDGPVMQKTMGWEASTERLYPRDGYLKEBIHQALXLDKGHYLVEFKTIYMAKKPVQLPGYYVDTKLDITSHNEDYTIVEQYERS EGRHHLFLYGMDELYK"			
ORIGIN			
Alignment Scores:			
Pred. No.:	4.93e-102	Length:	1431
Score:	1095.00	Matches:	201
Percent Similarity:	94.57%	Conservative:	8
Best Local Similarity:	90.95%	Mismatches:	12
Query Match:	90.20%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922a-12 (1-225) x AY678269 (1-1431)			
Qy	5	LysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGly	24
Db	16	GAGAGGTGATCAAGAAGTTTCATGCGCTTCAAGGTGCGCATGAGGGCTTCATGAAGGC	75
Qy	25	HisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGlyHisAsnThrVal	44
Db	76	CACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGACACCAAGCCGCGCC	135
Qy	45	LysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGln	64
Db	136	AACTGAAGGTGACCAAGGCGGCCCCCTGCGCTTGCGGACATCTGTCCCCCAG	195
Qy	65	PheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLys	84
Db	196	TTTCATGTAAGGCTCCAAAGCGGTACGTGAAGCAACCCCGCAGCATCCCCGATTACAAGAG	255
Qy	85	LeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyVal	104
Db	256	CTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAGAGACGGCGGTCTG	315
Qy	105	ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPhe	124
Db	316	GTGACCGTGAACCAAGACTCTCTCTGCAAGGACGCGCATCTTAACAAGGTGAAGATG	375
Qy	125	IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu	144
Db	376	CGCGGACCAACTTCCCCCGACGCGCCCGTAAATGCAAGAAGACCATGGCTGGAG	435
Qy	145	AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAla	164
Db	436	GCCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAGATCCACCAAGGCC	495
Qy	165	LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLys	184
Db	496	CTGAAGCTGAAGAGACGCGCGCCACATACCTGTGAGTTCAAGACCATCTACATGCGCAAG	555
Qy	185	LysProValGlnLeuProGlyTyrTyrTyrValAlaSerLysLeuAspIleThrSerHis	204
Db	556	AAAGCCGTGCAACTGCCCCGGCTACTACTACGTGACACCAAGCTGACATCACCTCCAC	615
Qy	205	AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHisLeuPhe	224
Db	616	AACGAGACTACACCATCGTGAACAGTACGAGCGCTCCGAGGCGCCACCACTGTTC	675
Qy	225	Leu 225	
Db	676	CTG 678	



RESULT 68	
AF272711	876 bp mRNA linear INV 26-SEP-2000
DEFINITION	Discosoma sp. SSAL-2000 red fluorescent protein (FP593) mRNA, complete cds.
ACCESSION	AF272711
VERSION	AF272711.1
KEYWORDS	GI:10304306
SOURCE	
ORGANISM	Discosoma sp. SSAL-2000
REFERENCE	Discosoma sp. SSAL-2000
AUTHORS	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.
TITLE	1 (bases 1 to 876)
JOURNAL	Fradekov, A.F., Chen, Y., Ding, L., Barbova, E.V., Matz, M.V. and Lukyanov, S.A.
PUBMED	Novel fluorescent protein from Discosoma coral and its mutants possesses a unique far-red fluorescence
REFERENCE	FEBS Lett. 479 (3), 127-130 (2000)
AUTHORS	10981720
TITLE	2 (bases 1 to 876)
JOURNAL	Fradekov, A.F., Chen, Y., Ding, L., Barbova, E.V., Matz, M.V. and Lukyanov, S.A.
PUBMED	Direct Submission
REFERENCE	Submitted (26-MAY-2000) Institute of Bioorganic Chemistry RAS, Mklukho-Makluya 16/10, Moscow 117871, Russia
AUTHORS	Location/Qualifiers
FEATURES	1. 876
SOURCE	/organism="Discosoma sp. SSAL-2000"
gene	/mol_type="mRNA"
CDS	/db_xref="taxon:137428"
	1. 876
	/gene="FP593"
	45. 737
	/gene="FP593"
	/note="d8FP593"
	/codon_start=1
	/product="red fluorescent protein"
	/protein_id="AAG16224.1"
	/db_xref="GI:10304307"
	/translation="MSCSKNVIKEMFRKVRMEGTVNGHEFEIKGEGRPVEGHCSV
	KLMTKGGPLPFAFDILSPQFGSKVYVKHPADLPDYKLSFPEGFKWERVNFEDG
	GVTVSQDSSLKDGCFIYEVKFIGVNFPSDGPVQRTRGWEASSRLYPRDVLKGL
	IHMALRLGGGHYLVFEEKSIYVKKPSVQLPGYVVYVSKLDMTSHNEDYTVVEQYEKTI
	QGRHHPFIKPLQ"
ORIGIN	
Alignment Scores:	2.53e-101
Pred. No.:	1085.50
Score:	95.58%
Percent Similarity:	88.05%
Best Local Similarity:	89.42%
Query Match:	2
DB:	2
US-10-006-922A-12 (1-225) x AF272711 (1-876)	
QY	1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
Db	45 ATGAGTTGTTCCAAAGATGTGATCAAGAGTTTCATGAGTTCAAGTTGATGGAAGGA 104
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGlyArgProTyrGluGly 40
Db	105 ACGGTCAATGGGACGAGTTTGAATAAAAGCGAAGGTGAAGGAGGCCCTTACGAAGGT 164
QY	41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
Db	165 CACTGTTCCTAAAGCTTATGTGTAACCAAGGCTGACCTTGCCATTGCTTTGATATT 224
QY	61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80
Db	225 TTGTACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAACACACCTGCCACATACCA 284
QY	81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100

Db 285 GACTATATAAAAGCTGTCAATTCCCTGAGGAGATTAAATGGAAAGGCTCATGAACCTTGAA 344

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 345 GACGGTGGCGTGGTTACTGTATCCCAAGATTCAGTTGAAAGACGGCTGTTTCATCTAC 404

Qy 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysThr 140

Db 405 GAGGTCAAGTTCATTGGGGTGAACCTTCTTGATGACCTGTATGACAGAGAGACA 464

Qy 141 MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160

Db 465 CGGGGCTGGGAAGCCAGCTCTGAGCGTTGTATCCTCGTATGGGGTGCTGAAGAGAGAC 524

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle 180

Db 525 ATCCATATGGCTCTGAGGCTGGAAGAGGGCGCCATTACCTCGTTGAATTCAAAAGTATT 584

Qy 181 TyrMetAlaLysLysPro--ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu 199

Db 585 TACATGTAAGAGAGCCTTCAGTGCAGTGGCCAGGCTACTATTATGTTGACTCCAAACTG 644

Qy 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGlyGly 219

Db 645 GATATGACGAGCCACAAGAGATTACACAGTCGTTGAGCAGTATGAAAAAACCCAGGGA 704

Qy 220 ArgHisHisLeuPheLeu 225

Db 705 CGCCACCATCCGTTCAATT 722

RESULT	69
AX686894	
LOCUS	AX686894
DEFINITION	Sequence 17 from Patent WO0127150.
ACCESSION	AX686894
VERSION	AX686894.1 GI:29409472
KEYWORDS	.
SOURCE	Discosoma sp.
ORGANISM	Discosoma sp. Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE	1 Lukyanov,S.A., Fradkov,A.F., Labas,Y.A., Matz,M.V. and Tersikh,A. Anthozoa derived chromo/fluoroproteins and methods for using the same Patent: WO 0127150-A 17 19-APR-2001; Clontech Laboratories Inc. (US) location/Qualifiers 1 . 876 /organism="Discosoma sp." /mol_type="unassigned DNA" /db_xref="taxon:86600"
JOURNAL	
AUTHORS	
TITLE	
FEATURES	
source	

Alignment Scores:	
Pred. No.:	2.53e-101
Score:	1085.50
Percent Similarity:	95.58%
Best Local Similarity:	88.05%
Query Match:	89.42%
DB:	6
	Length: 876
	Matches: 199
	Conservative: 17
	Mismatches: 9
	Indels: 1
	Gaps: 1

QY	1	M	e	a	r	s	e	r	S	e	r	L	y	A	s	n	V	a	i	i	e	L	y	S	G	L	P	H	e	m	e	t	A	r	g	P	h	e	L	y	S	v	a	L	A	r	g	M	e	t	G	L	I	G	L	I		20
Dd	45	A	T	G	A	C	T	T	G	T	T	C	C	A	G	A	T	G	T	G	A	C	A	G	A	G	T	T	C	A	T	G	A	G	T	T	C	A	T	G	A	G	T	T	C	G	T	A	T	G	A	A	G	A		104		
QY	21	T	H	R	V	A	L	S	N	G	L	I	H	I	S	G	L	P	H	E	G	L	I	L	E	G	L	I	G	L	I	G	L	I	A	R	G	P	R	O	T	T	R	G	L	I	G	L	I		40							
Dd	105	A	C	G	G	T	C	A	T	G	G	C	A	C	G	A	G	T	T	T	G	A	A	T	T	A	A	A	G	G	C	G	A	A	G	G	T	G	A	A	G	G	A	G	G	C	C	T	T	A	C	G	A	A	G	T		164



QY		41	HISAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db		165	CACGTGTTCCGTAAGCTTATGGTAACCAAGGGTGACCCTTGCATTGCTTTGATATT	224
QY		61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db		225	TTCATCACCACAATTTTCAGTATGGAAGCAAGGATATGTCAAACACCTGCCGACATACCA	284
QY		81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db		285	GACTATAAAAAGCTGTCAITTCCTGAGGGATTAAATGGGAAGGGTCATGAACTTTGAA	344
QY		101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db		345	GACGGTGGCGTGTTACTGTATCCCAAGATTCCAGTTTGAAAGACGGCTGTTTCATCTAC	404
QY		121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db		405	GAGGTCAAGTTTCATTGGGGTGAACCTTCTCTGATGACCTGTATGCAGAGAGAGACA	464
QY		141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db		465	CGGGGCTGGGAAGCCAGCTCTGAGCGTTTGATCTCGTGTGGGGTGCTGAAGAAGAGAC	524
QY		161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db		525	ATCATATGCGCTCTGAGGCTGGAAGAGGCGCCATTACTCGTTGAATTCAAAAAGTATT	584
QY		181	TyrMetAlaLysLysPro--ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeu	199
Db		585	TACATGTTAAGAAGCCCTTCAGTGCAGTTCAGGCTACTATTATGTTGACTCCAACCTG	644
QY		200	AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly	219
Db		645	GATATGACGAGCCACAACGAAGATTACACAGTCGTTGAGCAGTATGAAAAAACCCAGGGA	704
QY		220	ArgHisHisLeuPheLeu	225
Db		705	CGCCACCATCCGTTCAITT	722

```

RESULT 70
AF506027
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

```

CDS  
1..678  
/note="mRFP1; engineered variant of tetrameric red fluorescent protein from *Discosoma* deposited in GenBank  
Accession Number AF03369"

```

/codon_start=1
/translation_table=11
/product="monomeric red fluorescent protein"
/protein_id="AAM54544.1"
/db_xref="GI:21464838"
/translation="MASSEDAVKEPMRFKVRMEGSVNGHFEIEEGEGGRPYEGTQTA
KLKVTKGGLPLFAWDILSPQFYGSKAYVKGPAIDPYLKLSPFGKMERVMNPFEDG
GAVTVTLDSSLDQDEFIYKVKLRGTFNPSDGPVMQKTMGWBASTERMYPEDGALKGE
IMRLKLKDDGGHYDAEVKTTYMAKKPVQLPGAYKTIKLDITSHNEDYTIIVEQYRAE
GRHSTGA"

```

Alignment Scores:	
Pred. No.:	7.14e-95
Score:	1021.00
Percent Similarity:	89.14%
Best Local Similarity:	86.88%
Query Match:	84.10%
DB:	11
	Gaps: 0
	Length: 678
	Matches: 192
	Conservative: 5
	Mismatches: 24
	Indels: 0
	Gaps: 0

US-10-006-922A-12 (1-225) x AF506027 (1-678)

[illegible]



DEFINITION Synthetic construct monomeric DsRed variant gene for red fluorescent protein, partial cds.  
ACCESSION AB166761  
VERSION AB166761.1 GI:53748885  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 MatsuoKa,K.  
AUTHORS Monomeric DsRed variant with plant codon usage  
JOURNAL Unpublished  
TITLE 2 (bases 1 to 696)  
AUTHORS MatsuoKa,K.  
REFERENCES Direct Submision  
JOURNAL Submitted (08-MAR-2004) Ken MatsuoKa, RIKEN (The Institute of Physical and Chemical Research), Plant Science Center; 1-3-22 Suehirocho, Teurumi-ku, Yokohama 230-0045, Japan  
(E-mail:kenmat@psc.riken.jp,  
URL:http://mrg.psc.riken.go.jp/atrc/STRC%20lab%20index.htm,  
Tel:81-45-503-9575, Fax:81-45-503-9573)  
FEATURES  
source  
1..696  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
8..696  
/gene="monomeric DsRed variant"  
8..>696  
/gene="monomeric DsRed variant"  
/note="Synthetic monomeric DsRed protein with plant codon usage"  
/codon\_start=1  
/catal\_table=1  
/product="red fluorescent protein"  
/protein\_id="BAD52341.1"  
/db\_xref="GI:53748886"  
/translation="MASSEDDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEQTQTA  
KLKVTGGGFLPFAMDILSPQFGSKAYVKHPADIPDYKLSPBGFKMERVMNFEDG  
GVTVTQDSSLQDGBFIRYKVKLRGTFNPSDGPVMQKTMGEASTERYPEDGALKE  
IKMRKLKDGHYDAEVKTTYMAKKPVQLPGAYKTDIKLDTSHNEDYTIIVEQYERAE  
GRHSTGARPPES"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.37e-95 Length: 696  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x AB166761 (1-696)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 8 ATGCATCTTCCGAGGATGTGATTAAAGAGTTTATGAGATTCAAGGTGAGATGAGGGA 67  
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGlyGly 40  
DB 68 TCTGTTAATGACATGATGTCGAGATTGAGGGGAGAGAGAGAGACCTTACGAGGGG 127  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 128 ACTCAGACTGCTAAGCTGAAGGTTACTAAGGGGGGACCACTGCCTTCCCTGGGATATC 187  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 188 CTGCTCCACAATTCCAAATACGATCTAAGGCTTACGTCAAGCATCCAGCAGATATCCCA 247  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnGlu 100  
DB 248 GACTATTGGAAGCTCTCTTCCACAGAAAGATTCAAGTGGAGAGAGATTATGAACCTTCGAG 307

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyaspheIleTyr 120  
DB 308 GATGAGAGAGTTGTTACTGTACTCAAGATTCTCTCTCCCAAGATGGGAGTTCACTTAC 367  
QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 368 AAGGTGAAGCTCAGAGAACTAATTCCCATCTGATGACCAAGTATGCAAAAGAAGACT 427  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 428 ATGGGATGGGAGGCTTCTACTGAAGAATGTATCCAGAAATGGTGCTCTCAAGGAGAG 487  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 488 ATCAAGATGAGACTCAAGCTGAAGATGAGAGACATTATGATGCTGAAGTCAAGACTACC 547  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 548 TACATGGCTTAAGAACGACGTCACCTTCCAGAGCCTTACAAAGACTGACATCAAGCTCGAT 607  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 608 ATCACCTCCCATAAAGAGATTACCATCGTGAAACAATATGAGAGGGCGCTGAAGGTAGA 667  
QY 221 His 221  
DB 668 CAT 670  
RESULT 72  
DQ005472 4543 bp DNA circular SYN 07-MAY-2005  
LOCUS Cloning vector pSAT6A-RFP-N1, complete sequence.  
DEFINITION DQ005472  
ACCESSION DQ005472  
VERSION DQ005472.1 GI:63002541  
KEYWORDS  
SOURCE Cloning vector pSAT6A-RFP-N1  
ORGANISM Cloning vector pSAT6A-RFP-N1  
REFERENCE 1 (bases 1 to 4543)  
AUTHORS Chung,S.-M., Vyas,S., Kozlovskiy,S., Citovsky,V. and Tzfira,T.  
TITLE A versatile vector system for multiple gene expression in plants  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 4543)  
AUTHORS Chung,S.-M., Vyas,S., Kozlovskiy,S., Citovsky,V. and Tzfira,T.  
TITLE Direct Submision  
JOURNAL Submitted (12-APR-2005) Department of Biochemistry and Cell Biology, State University of New York at Stony Brook, Stony Brook, NY 11794, USA  
FEATURES  
source  
1..4543  
/organism="Cloning vector pSAT6A-RFP-N1"  
/mol\_type="other DNA"  
/db\_xref="taxon:326515"  
441..762  
/note="CamV 35S promoter"  
768..1089  
/note="CamV 35S promoter"  
1190..1320  
/note="translational enhancer 5'-UTR from tobacco etch virus"  
misc\_feature  
1321..1379  
/note="multiple cloning site; MCS"  
1380..2057  
/codon\_start=1  
/product="RFP"  
/protein\_id="AAY25379.1"  
/db\_xref="GI:63002542"  
/translation="MASSEDDVKEFMRFKVRMEGSVNGHEFEIEGEGRPYEQTQTA  
KLKVTGGGFLPFAMDILSPQFGSKAYVKHPADIPDYKLSPBGFKMERVMNFEDG  
GVTVTQDSSLQDGBFIRYKVKLRGTFNPSDGPVMQKTMGEASTERYPEDGALKE  
IKMRKLKDGHYDAEVKTTYMAKKPVQLPGAYKTDIKLDTSHNEDYTIIVEQYERAE  
GRHSTGA"







Db	1323	ATGGCTCTCTCCGAGGAGCTCATCAAGAGTTCATGCGCTTCAAGSTGCCATGGAGGGC	1382
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	1383	TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC	1442
Qy	41	HisAsnThrValIleuLeuValThrIleGlyGlyProLeuProPheAlaTrpAspIle	60
Db	1443	ACCCAGACCCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTTCGCTGGACATC	1502
Qy	61	LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro	80
Db	1503	CTGTCCCTCAGTTCAGTACGGCTCCAGAGCCTACGTGAAGCACCCCGCCGACATCCC	1562
Qy	81	AspTyrIleValLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu	100
Db	1563	GACTACTTGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	1622
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	1623	GACGGCGCGTGTGACCGTGACCCAGACTCTCCCTCGACGAGCGCGAGTTCACTTAC	1682
Qy	121	IleValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleIleThr	140
Db	1683	AAGGTGAAGCTGCGCGGCACCAACTTCCCTCCGACGGCCCGTAATGACAGAAGAACC	1742
Qy	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIleGlyIle	160
Db	1743	ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCCGAGGACGCGGCCCTGAAGGGCGAG	1802
Qy	161	IleHisIleValLeuIleValLeuIleValAspGlyGlyHisIleTyrLeuValGluPheIle	180
Db	1803	ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCCGAGGTCAAGACC	1862
Qy	181	TyrMetAlaIleValLeuProValGlnLeuProGlyTyrTyrTyrValAspSerIleVal	200
Db	1863	TACATGGCCAAGAGCCCGTGACAGCTGCCCGCGCCTACAAGACCGACATCAAGCTGAC	1922
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIleArg	220
Db	1923	ATCACTCCCAACGAGACTACACCATCGTGGAACAGTACGAGCGCGCCGAGGGCGGC	1982
Qy	221	His 221	
Db	1983	CAC 1985	
RESULT 74			
LOCUS	DQ005475	4558 bp	DNA
DEFINITION	Cloning vector pSAT6-RFP-N1, complete sequence.		
ACCESSION	DQ005475		
VERSION	DQ005475.1	GI:63002549	
KEYWORDS	Cloning vector pSAT6-RFP-N1		
SOURCE	Cloning vector pSAT6-RFP-N1		
ORGANISM	Cloning vector pSAT6-RFP-N1		
REFERENCE	other sequences; artificial sequences; vectors.		
AUTHORS	1 (bases 1 to 4558)		
TITLE	Chung,S.-M., Vyas,S. and Tzifira,T.		
JOURNAL	A versatile vector system for multiple gene expression in plants		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 4558)		
TITLE	Chung,S.-M., Vyas,S., Kozlovskiy,S., Citovsky,V. and Tzifira,T.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (12-APR-2005) Department of Biochemistry and Cell		
AUTHORS	Biology, State University of New York at Stony Brook, Stony Brook,		
TITLE	NY 11794, USA		
FEATURES	Location/Qualifiers		
Source	1..4558		
	/organism="Cloning vector pSAT6-RFP-N1"		
	/mol_type="other DNA"		
	/db_xref="taxon:326518"		
	441..762		
	/note="CamV 35S promoter"		

promoter	768..1089
enhancer	/note="CamV 35S promoter"
misc_feature	1190..1320
CDS	/note="translational enhancer 5'-UTR from tobacco etch virus"
	1321..1394
	/note="multiple cloning site; MCS"
	1395..2072
	/codon_start=1
	/product="RFP"
	/protein_id="AAV25384.1"
	/db_xref="GI:63002550"
	/translation="MASSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQTA
	KLKVTKGPLPFAMDILSPQOYGSKAYVKHPADI PDYLKLSFPEGFKWERVMNFE
	GVVTQDSLDQGEFIYKVKLRGTNFPSPDGPVMOKTMGWEASTERNYPEDGLKGE
	IKMRLLKDGHHYDAEVKTYMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYRAE
	GRHSTGA"
terminator	2079..2289
	/note="CamV 35S terminator"
CDS	complement(3498..4358)
	/note="amp resistance"
	/codon_start=1
	/product="Bla"
	/protein_id="AAV25385.1"
	/db_xref="GI:63002551"
	/translation="MSIOHFRVALIPFPAFCCLPVAHPETLVKVKDAEDQLGARVGY
	IEDLDNSKILESFRPERFRPMSTFKVLLCGAVLSRIDAGQEQGLRIHYSQNDLVE
	YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTITGPKELTAFILHMGDHYTRL
	DRWEPELNEAIPNDERDITWVPVAMATTIRKLITGELLTLASRQQLIDMEADKVA
	IRSAIPAGWFIADKSGAGERGSRGIIALGPDKPSRIIVITTTGSQATMDERNQLTA
	EIGASLIKHW"

ORIGIN

Alignment Scores:	7.29e-94	Length:	4558
Pred. No.:	1021.00	Matches:	192
Score:	89.14%	Conservative:	5
Percent Similarity:	86.88%	Mismatches:	24
Best Local Similarity:	84.10%	Indels:	0
Query Match:	11	Gaps:	0

US-10-006-922A-12 (1-225) x DQ005475 (1-4558)

Qy	1	MetArgSerSerIleAsnValIleIleGluPheMetArgPheIleValArgMetGluGly	20
Db	1395	ATGGCTCTCTCCGAGGAGCTCATCAAGAGTTCATGCGCTTCAAGGTGCCATGAGGGC	1454
Qy	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	1455	TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC	1514
Qy	41	HisAsnThrValIleuLeuIleValThrIleGlyGlyProLeuProPheAlaTrpAspIle	60
Db	1515	ACCCAGACCCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTTCGCTGGGACATC	1574
Qy	61	LeuSerProGlnPheGlnTyrGlySerIleValTyrValIleHisProAlaAspIlePro	80
Db	1575	CTGTCCCTCAGTTCAGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGCCGACATCCC	1634
Qy	81	AspTyrIleValLeuSerPheProGluGlyPheIleTyrGluArgValMetAsnPheGlu	100
Db	1635	GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTCGAG	1694
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	1695	GACGGCGCGTGTGACCGTGAACCAAGACTCTCTCCTCGACGAGCGCGAGTTCACTTAC	1754
Qy	121	IleValIlePheIleGlyValAsnPheProSerAspGlyProValMetGlnIleIleThr	140
Db	1755	AAGGTGAAGCTGCGCGGCACCAACTTCCCTCCGACGCGGCCCGTAATGACAGAAGACC	1814
Qy	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIleGlyGlu	160



Db 1815 ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCCGAGGACGGCCCTGAAGGGCGAG 1874  
Qy 161 ILeHIsLysAlAlLeuLySLeuLyAspRgLyGlyHIsTyRLeuValGluPhelysSerIle 180  
Db 1875 ATCAAGATGAGGCTGAAGCTGAAGACGGCGGCCACTACGACGCCGAGGTCAAGACCAACC 1934  
Qy 181 TYMeTAlAlLySbSProValGlnLeuProGlyTyTyTyTyRValAspSerLySLeuAsp 200  
Db 1935 TACATGGCCCAAGAAGCCCGTGACGCTGCCCGCGCCTACACAGACCGACATCAAGCTGGAC 1994  
Qy 201 ILeThrSerHIsAsnGluAspTyRThrIleValGlnGlnTyRGlUArgThrgLugLyArg 220  
Db 1995 ATCACTCCCAACAAGAGACTACCATCGTGAACAGTACGAGCGCGCGAGGGCGGC 2054  
Qy 221 HIs 221  
Db 2055 CAC 2057

RESULT 75  
AJ851287 5311 bp DNA circular SYN 12-NOV-2004  
LOCUS Cloning vector PRU1144 mRFP1 gene for mono Red fluorescent protein.  
DEFINITION AJ851287  
ACCESSION AJ851287 GI:55724880  
VERSION AJ851287.1 GI:55724880  
KEYWORDS mono Red fluorescent protein; mRFP1 gene.  
SOURCE Cloning vector PRU1144  
ORGANISM Cloning vector PRU1144  
REFERENCE 1 other sequences; artificial sequences; vectors.  
AUTHORS 1 Karunakaran, R. and Poole, P.S.  
TITLE High throughput promoter probe vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 5311)  
AUTHORS Poole, P.S.  
TITLE Direct Submission  
JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of  
READING, Whiteknights, Reading, RG6 6AJ, UNITED KINGDOM  
FEATURES  
source 1. 5311  
/organism="Cloning vector PRU1144"  
/mol\_type="other DNA"  
/db\_xref="taxon:299182"  
/lab\_host="Escherichia coli"  
326..1003  
/gene="mRFP1"  
326..1003  
/locus="mRFP1"  
/gene="mRFP1"  
/codon\_start=1  
/transl\_table=11  
/product="mono Red fluorescent protein"  
/protein\_id="CAH64892.1"  
/db\_xref="GI:55724881"  
/translation="MASSEDIYKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTA  
KLKVTXGGPLPFAMDIISPOFOYGSKAYVKHPADIPDYKLSPPEGFKMERVMNPEDG  
GVTVTODDSLQDGEPIYKYLKRGTFNPSDGPVMQKTMGWEASTERYMPEGALKE  
IKMRLKLDKGHYDAEVKTYTYMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAE  
GRHSTGA"

gene  
CDS  
326..1003  
/gene="mRFP1"  
326..1003  
/locus="mRFP1"  
/gene="mRFP1"  
/codon\_start=1  
/transl\_table=11  
/product="mono Red fluorescent protein"  
/protein\_id="CAH64892.1"  
/db\_xref="GI:55724881"  
/translation="MASSEDIYKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTA  
KLKVTXGGPLPFAMDIISPOFOYGSKAYVKHPADIPDYKLSPPEGFKMERVMNPEDG  
GVTVTODDSLQDGEPIYKYLKRGTFNPSDGPVMQKTMGWEASTERYMPEGALKE  
IKMRLKLDKGHYDAEVKTYTYMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERAE  
GRHSTGA"

ORIGIN

Alignment Scores:

Pred. No.:	8.78e-94	Length:	5311
Score:	1021.00	Matches:	192
Percent Similarity:	89.14%	Conservative:	5
Best Local Similarity:	86.88%	Mismatches:	24
Query Match:	84.10%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922A-12 (1-225) x AJ851287 (1-5311)

Qy 1 MetArgSerSerLySAsnValIleLySGluPhemeTArgPhelySValArgMetGluGly 20  
Db 326 ATGGCTCTCTCCAGAGACGTCAATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 385

Qy 21 ThrValAsnGlyHIsGluPheGlnIleGlnGlyGlnGlyArgProTyRGlUgLy 40  
Db 386 TCCGTGAACGGCCACAGAGTTCCAGATCGAGGGCGAGGGCGAGGCCCTTCAGAGGGC 445  
Qy 41 HIsAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60  
Db 446 ACCCAGACCGCCACAGCTGAAGGTGACCAAGGGCGGCCCTCGCCCTTCGCTGGACATC 505  
Qy 61 LeuSerProGlnPheGlnTyRgLySerLySValTyRValLySHIsProAlaAspIlePro 80  
Db 506 CTGTCCCTCAGTTCACAGTACGGCTCCAGAGGCTTACGTGAAGCACCCCGCGACATCCCC 565  
Qy 81 AspTyRlySbLeuSerPheProGlnGlyPhelySTrpGluArgValMetAsnPheliglu 100  
Db 566 GACTACTTGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 625  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyR 120  
Db 626 GACGCGCGCGTGTGACCGTGACCCAGACTCTCTCTGACAGACGGCGAGTTCAATCTAC 685  
Qy 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySbThr 140  
Db 686 AAGGTGAAGCTGCGCGGACCAACTTCCCTCCGACGGCCCGTAAATGCAAGAGAAGACC 745  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyRProArgAspGlyValLeuLySgLyGlu 160  
Db 746 ATGGGCTGGAGGCTTCCACCGAGCGGATGTACCCCGAGACGGCGCCCTGAAGGGCGAG 805  
Qy 161 ILeHIsLysAlAlLeuLySLeuLyAspRgLyHIsTyRLeuValGluPhelysSerIle 180  
Db 806 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCCACTACGACGCCGAGGTCAAGACCAACC 865  
Qy 181 TYMeTAlAlLySbSProValGlnLeuProGlyTyTyTyRValAspSerLySLeuAsp 200  
Db 866 TACATGGCCCAAGAAGCCCGTGACGCTGCCCGCGCCTACAAAGACGACATCAAGCTGGAC 925  
Qy 201 ILeThrSerHIsAsnGluAspTyRThrIleValGlnGlnTyRGlUArgThrgLugLyArg 220  
Db 926 ATCACTCCCAACAAGAGACTACCATCGTGAACAGTACGAGCGCGCGAGGGCGGC 985  
Qy 221 HIs 221  
Db 986 CAC 988

RESULT 76  
AJ851291 13079 bp DNA circular SYN 12-NOV-2004  
LOCUS Cloning vector PRU1161 gusA gene for beta-glucuronidase and mRFP1  
DEFINITION AJ851291  
ACCESSION AJ851291  
VERSION AJ851291.1 GI:55724891  
KEYWORDS beta-glucuronidase; gusA gene; mono Red fluorescent protein; mRFP1  
SOURCE Cloning vector PRU1161  
ORGANISM Cloning vector PRU1161  
REFERENCE 1 other sequences; artificial sequences; vectors.  
AUTHORS 1 Karunakaran, R. and Poole, P.S.  
TITLE High throughput promoter probe vectors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 13079)  
AUTHORS Poole, P.S.  
TITLE Direct Submission  
JOURNAL Submitted (21-OCT-2004) Poole P.S., School of AMS, University of  
READING, Whiteknights, Reading, RG6 6AJ, UNITED KINGDOM  
FEATURES  
source 1. 13079  
/organism="Cloning vector PRU1161"  
/mol\_type="other DNA"  
/db\_xref="taxon:299186"  
/lab\_host="Escherichia coli"  
186..863  
/gene="mRFP1"

gene



```
CDS
186. .863
/gene="mRFP1"
/codon_start=1
/transl_table=11
/product="mono Red fluorescent protein"
/protein_id="CAH64921.1"
/db_xref="GI:55724892"
/translation="MASSSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQTA
KLKVTKGGLPFAMDILSPQFGSKAYVKHPADIPDYKLSFPEGFKERVMNFDG
GVVTVTQDSSLQDGEFIYKVKLRGTNPSDGPVMQKTMGWEASTERYPEDGALKE
IKMRLKLDGSHYDAEVKTYMAKKPVQLPGAYYKTDIKLDTSHNEDYTIIVEQYERAE
GRHSTGA"
885. .2696
/gene="gusa"
885. .2696
/gene="gusa"
/codon_start=1
/transl_table=11
/product="beta-glucuronidase"
/protein_id="CAH64922.1"
/db_xref="GI:55724893"
/translation="MVRPVERTPREIKKLDGLMAFSLDRENGIDQRMWESALOESRA
IAYGSPNDQFADADIRNYAGVWYQREVFIPKGWAGQRIVLFPDVAATHYKXWVNNQ
EVMHQGGYTPFEADVTPYVITAGSKVRITVCNNELNMQRTIPGMVITDENGKKQSY
FHDFENYAGIHRSVMLYTPMTWVDITVTHVAQDCNHSVDWQVNAVANGSVLELRD
ADQVAVATGGTSGTLQVNPPLMQPGEGLYELCVTAKSQTECDIYPLRVGIRSVAY
KGEQFLINHKBPFYTFGRHEDADLRKGFDNLVMDHDLMDWIGANSYRTSHYPA
EEMLDWADEHGIVVIDETAAVGNLSLIGIPEAGNKPKEYSEBAVNGETQOAHLOAI
KELIARDKNHPSVVMMSIANEPTRRPGAREYFAPLAETRLDPTRPITCVNMFCD
AHTDTISDLFDVLCINRYGWYQSGDLETAEKVLEKELIMQEKLHQPITITEYGV
TLAHLHSMYTDMMSEBYQCAMLDMYHRAVFDVSAVVEGQVNNFADFATSGILRVGN
KKGIETRD RKPKSAFLLQKRWTMNFGKPKQGGKQ"
ORIGIN
```

Alignment Scores:  
Pred. No.: 2.64e-93 Length: 13079  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 11 Gaps: 0

```
US-10-006-922A-12 (1-225) x AJ851291 (1-13079)
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluLy 20
Db 186 ATGGCCTCCTCCGAGGACGTCAAGGAGTTCTCGCTTCAAGGTGCGCATGAGGGC 245
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugLy 40
Db 246 TCCGTGAACGGCCACGAGTTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 305
QY 41 HisAsnThrValyLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
Db 306 ACCCAGACCGCCAGAGTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGGACATC 365
QY 61 LeuSerProGlnPheGlnTyrgLySerLyValTyrgValLyHisProAlaAspIlePro 80
Db 366 CTGTCCCTCAGTTCAGTACGGGCTCCAGAGGCTACGTGAAGCAACCCGCGCATCCCC 425
QY 81 AspTyrgLyLeuSerPheProGluGlyPheLyTrpGluArgValMetAsnPheGlu 100
Db 426 GACTACTTGAAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTGAG 485
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCyaspheIleTy 120
Db 486 GACGGCGGCGTGTGACCGTGAACCCAGACTCTCTCCCTGACGAGCAGGGCGAGTTCTAC 545
QY 121 LysValLySpheIleGlyValAsnPheProSerAspGlyProValMetGlnLyLyThr 140
Db 546 AAGGTGAAGCTGCGCGCACCAACTTCCCTCCGACGGCCCGTAAATGCAAGAGAACC 605
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrgProArgAspGlyValLeuLySGlyGlu 160
```

```
Db 606 ATGGCCTGGAGGCGCTCCACCGAGCGGATGTACCCCGAGACGGCCCTGAAGGGCAG 665
QY 161 ILeHisLyAlaLeuLySLeuLyAspGlyGlyHisTyrgLeuValGluPheLySerIle 180
Db 666 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGCCACTACGACCGCGAGGTCAAGACACC 725
QY 181 TyrMetAlaLyLySProValGlnLeuProGlyTyrgTyrgTyrgValAspSerLySLeuAsp 200
Db 726 TACATGGCCCAAGAACCCGTCGAGCTGCCCGGCGCTTACAGACCGACATCAAGCTGAC 785
QY 201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgGluArgThrGluGlyArg 220
Db 786 ATCACTCCCAACAACGAGACTACCATCGTGAACAGTACGAGCGCGGAGGGCCGC 845
QY 221 His 221
Db 846 CAC 848
```

RESULT 77  
AY678271 678 bp DNA linear SYN 17-DEC-2004  
LOCUS  
DEFINITION Synthetic construct monomeric yellow fluorescent protein gene, complete cds.  
ACCESSION AY678271  
VERSION AY678271.1 GI:55420626  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

TITLE  
JOURNAL  
PUBMED  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
AUTHORS  
FEATURES  
source  
1. .678  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
1. .678  
/note="Honeydew; engineered variant of monomeric red fluorescent protein mRFP1 in GenBank Accession Number AF506027"  
/codon\_start=1  
/transl\_table=11  
/product="monomeric yellow fluorescent protein"  
/protein\_id="AAV52171.1"  
/db\_xref="GI:55420627"  
/translation="MASSSEDVIKEFMRFKVRMEGSVNGHEFEIEGEGRPYEGTQTA  
KLKVTKGGLPFAMDILSPQFGMSKAYVKHPADIPDYKLSFPEGFKERVMNFDG  
GVVTVTQDSSLQDGEFIYKVKLRGTNPSDGPVMQKTMGWEASTERYPEDGALKE  
IKMRLKLDGSHYDAEVKTYMAKKPVQLPGAYYKIDGKLDITSHNEDYTIIVEQYERAE  
GRHSTGA"

Alignment Scores:  
Pred. No.: 1.92e-93 Length: 678  
Score: 1007.00 Matches: 188  
Percent Similarity: 88.69% Conservative: 8  
Best Local Similarity: 85.07% Mismatches: 25  
Query Match: 82.95% Indels: 0  
DB: 11 Gaps: 0

```
US-10-006-922A-12 (1-225) x AY678271 (1-678)
```



Qy 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20  
||| |:::|  
Db 1 ATGGCCTCTCCGAGGACGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 60  
  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
:::|  
Db 61 TCCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120  
  
Qy 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
||| |  
Db 121 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGCTGGACATC 180  
  
Qy 61 LeuSerProGluPheGluTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
||| |:::|  
Db 181 CTGTCCCTCAGTTTCATGTGGGGCTCCAGAGCCTACGTGAAGCACCCCCGACATCCCC 240  
  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
||| |  
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 300  
  
Qy 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
||| |  
Db 301 GACGGCGCGGTGGTGACCGTGACCCAGACTCTCTCCCTGCAGAGCGCGAGTTTCATCTAC 360  
  
Qy 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGluLySLeuThr 140  
||| |  
Db 361 AAGGTGAAGCTGCGCGGCACCACTTCCCTCCGACGGCCCGTAATGCAGAAAGAGACC 420  
  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160  
||| |:::|  
Db 421 ATGGGCTGGCGGCCACCAACCGAGCGGATGTACCCCGAGACGGCGCCCTGAAGGGCGAG 480  
  
Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
||| |:::|  
Db 481 ATCAAGATGAGGCTGAAGCTGAAGACGGCGGCACTACGACGCCGAGGTCAAGACCACC 540  
  
Qy 181 TyrMetAlaLySLeuSProValGluLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
||| |:::|  
Db 541 TACATGGCCCAAGAGCCCGTGACGCTGCCCGGCCCTACAAGATTGACGGGAAGCTGGAC 600  
  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGluTyrGluArgThrGluGlyArg 220  
||| |  
Db 601 ATCACTTCCCAACAAGAGACTACCATCTGTGAACAGTACGAGCGCGCGGAGGGCGCG 660  
  
Qy 221 His 221  
|||  
Db 661 CAC 663  
  
RESULT 78  
AY678270 678 bp DNA linear SYN 17-DEC-2004  
LOCUS AY678270  
DEFINITION Synthetic construct monomeric orange-red fluorescent protein gene,  
complete cds.  
ACCESSION AY678270 GI:55420624  
VERSION AY678270.1  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
other sequences; artificial sequences.  
REFERENCE  
1 (bases 1 to 678)  
Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N.,  
Palmer,A.E. and Tsien,R.Y.  
Improved monomeric red, orange and yellow fluorescent proteins  
derived from Discosoma sp. red fluorescent protein  
Nat. Biotechnol. 22 (12), 1567-1572 (2004)  
15558047  
JOURNAL PUBMED  
2 (bases 1 to 678)  
REFERENCE  
AUTHORS Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and  
Tsien,R.Y.  
TITLE  
Direct Submission  
JOURNAL Submitted (06-JUL-2004) Pharmacology, University of California, San  
Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA

FEATURES  
Source  
Location/Qualifiers  
1..678  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
1..678  
/note="mTangerine; engineered variant of monomeric red  
fluorescent protein mRFP1 in GenBank Accession Number  
AF506027"  
/codon\_start=1  
/transl\_table=11  
/product="monomeric orange-red fluorescent protein"  
/protein\_id="AAV52170.1"  
/db\_xref="GI:55420625"  
/translation="MASSSEDVKEFMRFKVRMEGSVNGHEFIEGEGREPYEGTQTA  
KLKVTKGGLPEAMDILSPQFCYSKAYKHADIPLYKLSPPEGFKWERWNFEDG  
GVVTVTQDSSLQDGEFIYKVKLRGTINPBGDPVMOKTMGEASSERYPEDGALKGE  
IKMRLKLDGGHYDAEVKTTYMAKKPVQLPGAYKTDIKLDITSHNEDYTIIVLYERA  
GRHSTGA"  
  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.22e-93 Length: 678  
Score: 1002.00 Matches: 189  
Percent Similarity: 88.24% Conservative: 6  
Best Local Similarity: 85.52% Mismatches: 26  
Query Match: 82.54% Indels: 0  
DB: 11 Gaps: 0  
  
US-10-006-922A-12 (1-225) x AY678270 (1-678)  
  
Qy 1 MetArgSerSerLyAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20  
||| |:::|  
Db 1 ATGGCCTCTCCGAGGACGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 60  
  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
:::|  
Db 61 TCCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120  
  
Qy 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
||| |  
Db 121 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTGGCTGGACATC 180  
  
Qy 61 LeuSerProGluPheGluTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
||| |  
Db 181 CTGTCCCTCAGTTCTGTTAAGGCTCCAAAGGCTTACGTGAAGCACCCCGACATCCC 240  
  
Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
||| |  
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 300  
  
Qy 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
||| |  
Db 301 GACGGCGCGGTGGTGACCGTGACCCAGACTCTCTCCCTGCAGAGCGCGAGTTTCATCTAC 360  
  
Qy 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGluLySLeuThr 140  
||| |  
Db 361 AAGGTGAAGCTGCGCGGCACCAACTTCCCTCCGACGGCCCGTAATGCAAGAGAGACC 420  
  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160  
||| |:::|  
Db 421 ATGGGCTGGAGGCTCTCTCCGAGCGGATGTACCCCGAGGAGCGCGCCCTGAAGGGCGAG 480  
  
Qy 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
||| |  
Db 481 ATCAAGATGAGGCTGAAGCTGAAGAGCGCGGCACTTACGACGCCGAGGTCAAGACCACC 540  
  
Qy 181 TyrMetAlaLySLeuSProValGluLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
||| |  
Db 541 TACATGGCCCAAGAGCCCGTGACGCTGCCCGGCGCTACAAGACCGACATCAAGCTGGAC 600  
  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGluTyrGluArgThrGluGlyArg 220  
||| |



Db	601	ATCACCTCCCAACAGGAGTACACCAATCGTGAATTGTACGAGCGCGCCGAGGCGCCG	660
Qy	221	His 221	
Db	661	CAC 663	
RESULT	79		
LOCUS	AY679163	702 bp	DNA linear SYN 21-NOV-2004
DEFINITION	Synthetic construct mRFPmars-His6 gene, complete cds.		
ACCESSION	AY679163		
VERSION	AY679163.1	GI:55793413	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	other sequences; artificial sequences.		
AUTHORS	1 (bases 1 to 702)		
TITLE	Fischer,M., Haase,I., Simmeth,E., Gerisch,G. and Muller-Taubenberger,A.		
JOURNAL	A brilliant monomeric red fluorescent protein to visualize cytoskeleton dynamics in Dictyostelium		
PUBMED	FEBS Lett. 577 (1-2), 227-232 (2004)		
15527790			
2 (bases 1 to 702)			
Fischer,M.J. and Mueller-Taubenberger,A.			
Direct Submission			
Submitted (07-JUL-2004)	Chemistry, Technical University Munich, Lichtenbergstr. 4, Garching 85747, Germany		
Location/Qualifiers			
1. .702	/organism="synthetic construct"		
/mol_type="other DNA"			
/db_xref="taxon:32630"			
1. .702	/note="red fluorescent protein"		
/codon_start=1			
/transl_table=11			
/product="mRFPmars-His6"			
/protein_id="AAV65051.1"			
/db_xref="GI:55775541"			
/translation="MASSSEDVKEFMRFKYKNQGSVNGHEFEIEGEGEGRPYEGTQTA			
KLKVTKGSPLEPRAMDILSPQFQYGSKAYVKHPADIPDYKLSPFGFKWERIMNFEDG			
GVVTVTQDSLTLDGSEFIYKVKLRGTNFPSPDGPVQKKTGWEASTERYPEDGALKGE			
IKMRLLKLDGGHYDAEVKTTYKAKKQVQLPGAYKTDIKLDIISHNEDYTIIVEQYERAE			
GRHSTGAGSHHHHHH"			
676. .681			
/note="Region: linker"			
682. .699			
/note="Region: histidine-tag; 6xHis"			
ORIGIN			
Alignment Scores:			
Pred. No.:	4.26e-92	Length:	702
Score:	994.00	Matches:	186
Percent Similarity:	87.78%	Conservative:	8
Best Local Similarity:	84.16%	Mismatches:	27
Query Match:	81.88%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922A-12 (1-225) x	AY679163 (1-702)		
Qy	1	MeArGSeRSeRlySAsnValIllelySGluPheMeArGpHeLyVaLaRGMeTglUgLy	20
Db	1	ATGGCATCATCAGAAGATGTTATTAAAGAAATTATGAGATTCAAAGTTAAATGGAAGGT	60
Qy	21	ThrValaNGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrgLugLy	40
Db	61	TCAGTTAACCGTCATGATTTGAAATGGAAGGTGAAGGTGAAGGTGACCATATGAAGGT	120
Qy	41	HisAenThVallySleuLySvalThrySGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	ACACAACAGCAGCAAAATTAAAGTTACAAAAGTGTCATTCATTCATTGCATGGGATATT	180

Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTTCAACCACAATTTCATATATGGTAGTAAGCATATGTAAACATCCAGACAGATATTTCCA	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GATTATCTTAAACTTTCATTTCCAGAAGGTTTAAATGGGAAGAATTATGAATTTTGAA	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GATGGTGTGTGTGTACAGTTACTTCAAGATTTCACATTACAGATGTGGAATTATTTAT	360
Qy	121	LysValLysPheIleGlyValAsnDheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAAGTTAAATTAAAGAGTACCATTTCATCCATCAGATGGTCCAGTAATGCATAAAAACT	420
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGTGGGAAGCATCAACAGAAAGATGTATCCAGAAGATGGTCACTAAAAAGGTGAA	480
Qy	161	IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATTAAATGAGATTAAAACTTAAAGATGGTGTCTATTATGATGCAGAACTTAAACAAC	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TATAAAGCTAAACAAACAAGTTCATTATCCAGGTCTTATATAAACAGATATTAAATTAGAT	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATTATTTCACATATGAAGATTATACCTATGTGTGAACAATATGAAGACAGCAAGGTAGA	660
Qy	221	His 221	
Db	661	CAT 663	
RESULT 80			
LOCUS	AY678265	711 bp	DNA linear SYN 17-DEC-2004
DEFINITION			Synthetic construct monomeric orange fluorescent protein gene, complete cds.
ACCESSION	AY678265		
VERSION	AY678265.1	GI:55420614	
KEYWORDS			
SOURCE			synthetic construct
ORGANISM			synthetic construct
REFERENCE			other sequences; artificial sequences.
AUTHORS			1 (bases 1 to 711)
TITLE			Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N., Palmer,A.E. and Tsien,R.Y.
JOURNAL			Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein
PUBMED			Nat. Biotechnol. 22 (12), 1567-1572 (2004)
REFERENCE			15558047
AUTHORS			2 (bases 1 to 711)
TITLE			Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and Tsien,R.Y.
JOURNAL			Direct Submission
PUBMED			Submitted (06-JUL-2004) Pharmacology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
REFERENCE			Location/Qualifiers
AUTHORS			1..711
TITLE			/organism="synthetic construct"
JOURNAL			/mol_type="other DNA"
PUBMED			/db_xref="taxon:32630"
REFERENCE			1..711
AUTHORS			/note="mOrange; engineered variant of monomeric red fluorescent protein mRFP1 in GenBank Accession Number AF506027; N- and C-termini replaced with equivalent residues from EGFP"
JOURNAL			/codon_start=1
PUBMED			/transl_table=11
REFERENCE			/product="monomeric orange fluorescent protein"
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			



/protein\_id="AAV52165.1"  
/db\_xref="GI:55420615"  
/translation="MWSKGEEDNMAIIKEFMRFKVMESVNGHEFIEEGEGRPYE  
GEQTAKLKVTXGGPLPFAWDILSPQFTYGSKAYVKHPADI PDYFKLSPEGEKMERVM  
NFEDGVVTVTQDSSLQDGEFIYKVKLRGTNFPSPDPMQKTMGEASSERMYPEDG  
ALKEIKMRLKLDKGGHYTSVKTYYKAKKPVQLPGAIVIGIKLDITSHNEDYTIVEQ  
YERAEGRHSTGMDLYK"

ORIGIN

Alignment Scores:				
Pred. No.:	1.4e-91	Length:	711	
Score:	989.00	Matches:	185	
Percent Similarity:	88.37%	Conservative:	5	
Best Local Similarity:	86.05%	Mismatches:	25	
Query Match:	81.47%	Indels:	0	
DB:	11	Gaps:	0	

US-10-006-922A-12 (1-225) x AY678265 (1-711)

QY	7	ValIleIleYsGluPheMetArgPheIysValArgMetGluGlyThrValAsnGlyHisGlu	26
DB	34	ATCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGGCTCCGTGAACGGCCACGAG	93
QY	27	PheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsnThrValIysLeu	46
DB	94	TTCCAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGCTTTCAGACCGCTTAAGCTG	153
QY	47	LysValThrIysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln	66
DB	154	AAGGTGACCAAGGTTGCCCTGCTTCGCTGGACATCTGCTCCCTCAGTTCAAC	213
QY	67	TyrGlySerIysValTyrValIysHisProAlaAspIleProAspTyrIysIysLeuSer	86
DB	214	TACGGCTCCAAGGCTTACGTGAAGACCCCGCGACATCCCGCACTACTTCAAGCTGTCC	273
QY	87	PheProGluGlyPheIysTrpGluArgValMetAsnPheGluAspGlyValValThr	106
DB	274	TTCCCGGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAGGACGGCGGTGTGACC	333
QY	107	ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIysValIysPheIleGly	126
DB	334	GTGACCAAGGACTCTCCCTGACAGGCGCGAGTTTCATCTACAGGTGAAGCTGCGCGGC	393
QY	127	ValAsnPheProSerAspGlyProValMetGlnIysIysThrMetGlyTrpGluAlaSer	146
DB	394	ACCAACTTCCCTCCGACGGCCCGTATGCAGAAGAAGACCAATGGGCTGGAGGCTTCC	453
QY	147	ThrGluArgLeuTyrProArgAspGlyValIleuIysGlyIleHisIysAlaLeuIys	166
DB	454	TCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGATGAGGCTGAAG	513
QY	167	LeuIysAspGlyGlyHisIysTyrLeuValGluPheIysSerIleTyrMetAlaIysIysPro	186
DB	514	CTGAAGGACGGCGGCCACTACACCTCCGAGGTCAAGACCACTTCAAGGCCAAGAGCCCC	573
QY	187	ValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAspIleThrSerHisAsnGly	206
DB	574	GTGCAAGTGGCCCGGCGCTTACATGTCGGCATCAAGTTGGACATCACTCCCAACAGAG	633
QY	207	AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis	221
DB	634	GACTACACCATCGTGAACAGTACGAACGCGCGGAGGCGGCCAC	678

RESULT 81  
AY678264 711 bp DNA linear SYN 17-DEC-2004  
LOCUS AY678264  
DEFINITION Synthetic construct monomeric red fluorescent protein gene,  
complete cds.  
ACCESSION AY678264  
VERSION AY678264.1 GI:55420612  
KEYWORDS  
SOURCE .  
. ORGANISM synthetic construct  
synthetic construct

other sequences; artificial sequences.

REFERENCE	1 (bases 1 to 711)
AUTHORS	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Giepmans,B.N., Palmer,A.E. and Tsien,R.Y.
TITLE	Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein
JOURNAL	Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED	15558047
REFERENCE	2 (bases 1 to 711)
AUTHORS	Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and Tsien,R.Y.
TITLE	Direct Submision
JOURNAL	Submitted (06-JUL-2004) Pharmacology, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
FEATURES	Location/Qualifiers
source	1. .711 /organism="synthetic construct" /mol_type="other DNA" /db_xref="taxon:32630"

CDS

1. .711  
/note="mCherry: engineered variant of monomeric red  
fluorescent protein mRFP1 in GenBank Accession Number  
AF506027; N- and C-terminal replaced with equivalent  
residues from EGFP"  
/codon\_start=1  
/transl\_table=11  
/product="monomeric red fluorescent protein"  
/protein\_id="AAV52164.1"  
/db\_xref="GI:55420613"  
/translation="MWSKGEEDNMAIIKEFMRFKVMESVNGHEFIEEGEGRPYE  
GTQTAKLKVTXGGPLPFAWDILSPQFTYGSKAYVKHPADI PDYFKLSPEGEKMERVM  
NFEDGVVTVTQDSSLQDGEFIYKVKLRGTNFPSPDPMQKTMGEASSERMYPEDG  
ALKEIKMRLKLDKGGHYDAEVKTTYKAKKPVQLPGAIVNINIKLDITSHNEDYTIVEQ  
YERAEGRHSTGMDLYK"

ORIGIN

Alignment Scores:				
Pred. No.:	2.24e-91	Length:	711	
Score:	987.00	Matches:	184	
Percent Similarity:	88.84%	Conservative:	7	
Best Local Similarity:	85.58%	Mismatches:	24	
Query Match:	81.30%	Indels:	0	
DB:	11	Gaps:	0	

US-10-006-922A-12 (1-225) x AY678264 (1-711)

QY	7	ValIleIleYsGluPheMetArgPheIysValArgMetGluGlyThrValAsnGlyHisGlu	26
DB	34	ATCATCAAGGAGTTTCATGCGCTTCAAGGTGCAATGAGGGCTCCGTGAACGGCCACGAG	93
QY	27	PheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsnThrValIysLeu	46
DB	94	TTCCAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGCAACCGAGCGCCAAAGCTG	153
QY	47	LysValThrIysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln	66
DB	154	AAGGTGACCAAGGTTGCCCTGCTTCGCTGGACATCTGTCCTCAGTTCAATG	213
QY	67	TyrGlySerIysValTyrValIysHisProAlaAspIleProAspTyrIysIysLeuSer	86
DB	214	TACGGCTCCAAGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAGGACGGCGGTGTGACC	273
QY	87	PheProGluGlyPheIysTrpGluArgValMetAsnPheGluAspGlyValValThr	106
DB	274	TTCCCGGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAGGACGGCGGTGTGACC	333
QY	107	ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIysValIysPheIleGly	126
DB	334	GTGACCAAGGACTCTCCCTGACAGGCGCGAGTTTCATCTACAAAGTGAAGCTGCGCGGC	393
QY	127	ValAsnPheProSerAspGlyProValMetGlnIysIysThrMetGlyTrpGluAlaSer	146
DB	394	ACCAACTTCCCTCCGACGGCCCGTATGCAAGAAGACCATGGGCTGGAGGCTTCC	453



Qy 147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 454 TCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGGAGATCAAGCAGAGGCTGAAG 513  
Qy 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 514 CTGAAGACGGCGGCCACTACGAGCTGAGTCAAGACCACCTTACAGGCCAAGAGAGCC 573  
Qy 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 574 GTGAGCTGCGCGCGCTTACAACGTCAACATCAAGTTGACATCACTCCCAACAGAG 633  
Qy 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 634 GACTACCATCGTGGAGACAGTACGAACGGCGGAGGCCGCCAC 678  
RESULT 82  
AY786536 681 bp mRNA linear INV 01-DEC-2004  
LOCUS Discosoma sp. LW-2004 fluorescent protein raspberry mRNA, complete  
DEFINITION cds.  
ACCESSION AY786536  
VERSION AY786536.1 GI:55792806  
KEYWORDS  
SOURCE Discosoma sp. LW-2004  
ORGANISM Discosoma sp. LW-2004  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
TITLE Evolution of new nonantibody proteins via iterative somatic  
JOURNAL hypermutation  
PUBMED Proc. Natl. Acad. Sci. U.S.A. 101 (48), 16745-16749 (2004)  
1556995  
REFERENCE 2 (bases 1 to 681)  
AUTHORS Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
TITLE Direct Submission  
JOURNAL Submitted (18-OCT-2004) Department of Pharmacology, University of  
California-San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA  
FEATURES  
source Location/Qualifiers  
1. .681  
/organism="Discosoma sp. LW-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:301246"  
1. .681  
/codon\_start=1  
/product="fluorescent protein raspberry"  
/protein\_id="AAV65486.1"  
/db\_xref="GI:55792807"  
/translation="MVSKEEVIKEMRFKVRMEGSVNGHEPEIEGEGEPYEQTQT  
AKLKTGKGPLPFAMDILSPQCMYSGKYVKPADIPDYLKLSPEGFKMERVMNFED  
GGVVTVDSSLDGEFIYKVKLRGTNPPSDGPVMQKTKTWEASSESRMYPEDGALKG  
EMKRLKLDGSHYDAEVKTTYMAKPEVQLPGA YKTDIKDITSHNEDYTIIVEQYERA  
EGRHSTGA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.41e-91 Length: 681  
Score: 985.00 Matches: 185  
Percent Similarity: 88.02% Conservative: 6  
Best Local Similarity: 85.25% Mismatches: 26  
Query Match: 81.14% Indels: 0  
DB: 2 Gaps: 0  
US-10-006-922A-12 (1-225) x AY786536 (1-681)  
Qy 5 LysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGly 24  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 16 GAGGAGGTCATCAAGAGTTCATCGCTTCAAGGTGCGCATGAGGCTCCGTGAACGGC 75  
Qy 25 HisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsnThrVal 44  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 76 CACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGCACCAGACGGCC 135  
Qy 45 LysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGln 64  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 136 AAGCTGAAGGTGACCAAGGGTGGCCCCCTTGCCCTTGCGGAGCATCTGTCCCTCAG 195  
Qy 65 PheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLys 84  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 196 TGCATGTACGGCTCCAAGGGCTTACGTGAAGCACCCCGGCACATCCCGACTACTGAAG 255  
Qy 85 LeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyVal 104  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 256 CTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTTCGAGGACGGCGCGCTG 315  
Qy 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPhe 124  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 316 GTGACCGTGAACCCAGACTCTCTCCCTGCAGAGACGGCGAGTTTCAACAAGTGAAGCTG 375  
Qy 125 IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu 144  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 376 CGCGGACCAACTTCCCTCCGACGGCCCCGTATATGACAGAAGAACATGGCTGGAG 435  
Qy 145 AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAla 164  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 436 GCCTCTCCGAGCGGATGTATCCCGACGAGACGGCGCCCTGAAGGGCGAGATGAAGATGAGG 495  
Qy 165 LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLys 184  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 496 CTGAAGCTGAAGACGCGCGCCACTTACGACGCCGAGGTCAAGACCACTTACATGCGCAG 555  
Qy 185 LysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHis 204  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 556 AAGCCCGTCAAGCTGCGCGCGCTTACAAGACCGACATCAAGTTGACATCACTCCAC 615  
Qy 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||  
Db 616 AACGAGACTACACCATCGTGGAGACGTACGAGCGCGGAGGGCGGCCAC 666  
RESULT 83  
AY678266 711 bp DNA linear SYN 17-DEC-2004  
LOCUS Synthetic construct monomeric red-orange fluorescent protein gene,  
DEFINITION complete cds.  
ACCESSION AY678266  
VERSION AY678266.1 GI:55420616  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE other sequences; artificial sequences.  
AUTHORS 1 (bases 1 to 711)  
Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.,  
Palmer, A.E. and Tsien, R.Y.  
TITLE Improved monomeric red, orange and yellow fluorescent proteins  
JOURNAL derived from Discosoma sp. red fluorescent protein  
PUBMED Nat. Biotechnol. 22 (12), 1567-1572 (2004)  
1558047  
REFERENCE 2 (bases 1 to 711)  
AUTHORS Shaner, N.C., Campbell, R.B., Steinbach, P.A., Palmer, A.E. and  
Tsien, R.Y.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-2004) Pharmacology, University of California, San  
Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA  
FEATURES  
source Location/Qualifiers  
1. .711  
/organism="synthetic construct"  
/mol\_type="other DNA"  
/db\_xref="taxon:32630"  
1. .711  
/note="msStrawberry; engineered variant of monomeric red  
fluorescent protein mRFP1 in GenBank Accession Number  
AF506027; N- and C-terminal replaced with equivalent  
residues from EGFP"  
/codon\_start=1



/transl\_table=11  
/product="monomeric red-orange fluorescent protein"  
/protein\_id="AAV52166.1"  
/db\_xref="GI:55420617"  
/translation="MVSKEGVNMAIIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYE  
GTQAKLVTKGGPLPFAMWDLITENFTYGSKAYVKHPADIPDYLKLSFPEGFKWERVM  
NFEDEGVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKTMGWEASSERMYPEDG  
ALKEIKMRLKLDGGHYDAEVKTYKAKKPVQLPGAYIVGIIKLDITSHNEDYTIIVEL  
YERAEGRHSTGMDELYK"

ORIGIN

Alignment Scores:	6.03e-90	Length:	711
Pred. No.:	973.00	Matches:	182
Score:	87.44%	Conservative:	6
Percent Similarity:	87.44%	Mismatches:	27
Best Local Similarity:	84.65%	Indels:	0
Query Match:	80.15%	Gaps:	0
DB:	11		

US-10-006-922A-12 (1-225) x AY678266 (1-711)

Qy	7	ValIleLySGluPheMetArgPheLySValArgMetGluGlyThrValAsnGlyHisGlu	26
Db	34	ATCATCAAGAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGGCCACGAG	93
Qy	27	PheGluIleGluGlyGluGlyArgProTyrGlyGlyHisAsnThrValLySLeu	46
Db	94	TTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGCACCCAGACCGCCAAGCTG	153
Qy	47	LySValThrLySGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln	66
Db	154	AAGGTGACCAAGGTTGGCCCCCTGCCCTTGCCTGGACATCTTAACCCCACTTCAACC	213
Qy	67	TyrGlySerLySValTyrValLySHisProAlaAspIleProAspTyrLySlySLeuSer	86
Db	214	TACGGCTCCAAAGCCTTACGTGAAGCACCCCGCGACATCCCGACTACTTGAAGCTGTCC	273
Qy	87	PheProGluGlyPheLySTrpGluArgValMetAsnPheGluAspGlyGlyValValThr	106
Db	274	TTCCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAGGACGGCGCGGTGAC	333
Qy	107	ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLySValLySPheIleGly	126
Db	334	GTGACCCAGACTCTCCCTGCAGACGGCGAGTTTCATCTACAAGGTGAAGCTGCGCGGC	393
Qy	127	ValAsnPheProSerAspGlyProValMetGlnLySlySThrMetGlyTyrGluAlaSer	146
Db	394	ACCAACTTCCCTCCGACGGCCCCGTATGCAAGAAGAACCACCTGCGGCTGGGAGCGCTCC	453
Qy	147	ThrGluArgLeuTyrProArgAspGlyValLeuLySGlyIleHisLySAlaLeuLyS	166
Db	454	TCCGAGCGGATGTACCCCGAGGACGGCGCTGAAGGGCGAGATCAAGATGAGGCTGAG	513
Qy	167	LeuLyAspGlyGlyHisTyrLeuValGluPheLySerIleTyrMetAlaLySlySPro	186
Db	514	CTGAAGGACGGCGGCCCACTACGACGCTGAGGTCAAGACCACTTACAAGGCCAAGAAGCCC	573
Qy	187	ValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAspIleThrSerHisAsnGlu	206
Db	574	GTGCAGCTGCCCGCGCCTACATCGTCGCATCAAGTTGACATCACTCCCAACAGAG	633
Qy	207	AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis	221
Db	634	GACTACACCATCGTGAAGTGTACGAACGCGCGAGGGCGGCCAC	678

RESULT 84  
AY786537  
LOCUS AY786537 681 bp mRNA linear INV 01-DEC-2004  
DEFINITION Discosoma sp. LW-2004 fluorescent protein plasm mRNA, complete cds.  
ACCESSION AY786537  
VERSION AY786537.1 GI:55792808  
KEYWORDS  
SOURCE Discosoma sp. LW-2004

ORGANISM Discosoma sp. LW-2004  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;  
Corallimorpharia; Discosomatidae; Discosoma.  
REFERENCE  
AUTHORS Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
TITLE Evolution of new nonantibody proteins via iterative somatic  
hypermutation  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (48), 16745-16749 (2004)  
PUBMED 15556995  
REFERENCE 2 (bases 1 to 681)  
AUTHORS Wang, L., Jackson, W.C., Steinbach, P.A. and Tsien, R.Y.  
TITLE Direct Submission  
JOURNAL Submitted (18-OCT-2004) Department of Pharmacology, University of  
California-San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA  
FEATURES  
source location/Qualifiers  
1..681  
/organism="Discosoma sp. LW-2004"  
/mol\_type="mRNA"  
/db\_xref="taxon:301246"  
1..681  
/codon\_start=1  
/product="Fluorescent protein plasm"  
/protein\_id="AAV65487.1"  
/db\_xref="GI:55792809"  
/translation="MVSKEGVNMAIIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEQTQ  
ARLVTKGGPLPFAMWDLISPOIMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFE  
DGVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKTMGWEASSERMYPEDGAL  
EMKMRRLKLDGGHYDAEVKTYKAKKPVQLPGAYKTDIKLDITSHNEDYTIIVEQYERA  
EGRHSTGA"

CDS

Alignment Scores:	7.24e-90	Length:	681
Pred. No.:	972.00	Matches:	181
Score:	87.10%	Conservative:	8
Percent Similarity:	87.10%	Mismatches:	28
Best Local Similarity:	83.41%	Indels:	0
Query Match:	80.07%	Gaps:	0
DB:	2		

US-10-006-922A-12 (1-225) x AY786537 (1-681)

Qy	5	LySAsnValIleLySGluPheMetArgPheLySValArgMetGluGlyThrValAsnGly	24
Db	16	GAGAGGTTCATCAAGAGTTTCATGCGCTTCAAGGAGCACATGAGGGCTCCGTGAACGGC	75
Qy	25	HisGluPheGluIleGluGlyGlyGlyArgProTyrGlyGlyHisAsnThrVal	44
Db	76	CACGAGTTCGAGATCGAGGCGAGGGCGCGCCCTTACGAGGACCCAGACCCAGACCGCC	135
Qy	45	LySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGln	64
Db	136	AGGTGAAGGTGACCAAGGTTGGCCCCCTGCCCTTGGCGCATCTGTCCCTCAG	195
Qy	65	PheGlnTyrGlySerLySValTyrValLySHisProAlaAspIleProAspTyrLySlyS	84
Db	196	ATCATGTACGGCTCCCAAGGCTTACGTGAAGCACCCCGCGACATCCCGACTACTGAAG	255
Qy	85	LeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGluAspGlyGlyVal	104
Db	256	CTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGGTGATGAACCTTCGAGGACGGCGCGGTG	315
Qy	105	ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLySValLyS	124
Db	316	GTGACCGTGACCCAGACTCTCCCTGCAGACGGCGAGTTTCATCTACAAGGTGAAGGTG	375
Qy	125	IleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThrMetGlyTyrGlu	144
Db	376	CGCGGACCAACTTCCCTCCGACGGCCCCGTATGCAAGAAGACCATGGGCTGGAG	435
Qy	145	AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyIleHisLySAla	164
Db	436	GCCTCTCCGAGCGGAGTGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATGAAGTAGG	495



```

Qy      165 LeuLYsLeuLYsAspGLyGLyHisIstYrLeuValGLuPhelySerIleTyrMetalAlaYs 184
      |||:::|||||
Db      496 CTGAGGCTGAAGACGCGCGCCACTACGACGCCGAGGTCAAGACCACCTACATGCGCAAG 555

Qy      185 LysProValGlnLeuProGLyTyrTyrTyrValAspSerLYsLeuAspIleThrSerHis 204
      |||
Db      556 AAGCCCGTGAAGCTGCCCCGCGCCTACAGACCACATCAAGCTGACATCACTCCAC 615

Qy      205 AsnGLuAspTyrThrIleValGLuGlnTyrGLuArgThrGLuGLyArgHis 221
      |||
Db      616 AACGAGACTACACCATCTGTGGAACAGTACGAGCGCGGAGGCGCCAC 666

RESULT 85
AY678267          711 bp      DNA      linear      SYN 17-DEC-2004
LOCUS             Synthetic construct monomeric yellow-orange fluorescent protein
DEFINITION        gene, complete cds.
ACCESSION         AY678267
VERSION           AY678267.1  GI:55420620
KEYWORDS          .
SOURCE            synthetic construct
ORGANISM          synthetic construct
REFERENCE         1 (bases 1 to 711)
AUTHORS           Shaner,N.C., Campbell,R.E., Steinbach,P.A., Gilemans,B.N.,
                  Palmer,A.E. and Tsien,R.Y.
TITLE             Improved monomeric red, orange and yellow fluorescent proteins
                  derived from Drosophila sp. red fluorescent protein
JOURNAL           Nat. Biotechnol. 22 (12), 1567-1572 (2004)
PUBMED            15558047
REFERENCE         2 (bases 1 to 711)
AUTHORS           Shaner,N.C., Campbell,R.E., Steinbach,P.A., Palmer,A.E. and
                  Tsien,R.Y.
TITLE             Direct Submission
JOURNAL           Submitted (06-JUL-2004) Pharmacology, University of California, San
                  Diego, 9500 Gilman Dr., La Jolla, CA 92093-0647, USA
FEATURES          Location/Qualifiers
          source   1..711
                   /organism="synthetic construct"
                   /mol_type="other DNA"
                   /db_xref="taxon:32630"
          CDS      1..711
                   /note="mBanana; engineered variant of monomeric red
                   fluorescent protein mRFP1 in GenBank Accession Number
                   AF506027; N- and C-termini replaced with equivalent
                   residues from EGFP"
                   /codon_start=1
                   /transl_table=11
                   /product="monomeric yellow-orange fluorescent protein"
                   /protein_id="AAV52167.1"
                   /db_xref="GI:55420621"
                   /translation="MVSKEENMNAVIKEMRKYRMESVNGHEFIDEGEGRPYE
                   GTOTAKLKVTKGGPLPFAWDILSPQFCYGSKAYVKHPTGIPDYFKLSFPEGFKMERVM
                   NPEDGSVTVTAQDSLODBEFIVKLRGTNPSDGPVMQKTMGEASSEMYPEDG
                   ALKGEIKMRLKLDGSHYSABTKTTYKAKKPVQDLPAYIAGEKIDITSHNEDYTIIVEL
                   YERAEGRHSTGMDELYK"

ORIGIN
Alignment Scores:
Pred. No.:      1.28e-88      Length:      711
Score:          960.00        Matches:      180
Percent Similarity: 86.05%      Conservative: 5
Best Local Similarity: 83.72%      Mismatches: 30
Query Match:    79.08%        Indels:      0
DB:             11           Gaps:        0

US-10-006-922A-12 (1-225) x AY678267 (1-711)

Qy      7 ValIleLYsGLuPhMetArgPheLYsValArgMetGLuLYThrValAsnGLyHisGLu 26
      |||
Db      34 GTCATCAAGAGATTGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGCCACGAG 93

Qy      27 PheGLuIleGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLuGLu 46

```

D	b		94	TTCCGATCGAGGGCGAGGGCGGGCCCTTACGAGGGCACCCAGACCGCCAAGCTG	153
O	y		47	LysValThrIySGlyGIYProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln	66
D	b		154	AAGGTGACC AAGGGTGCCCCCTGCCTTCGCTGGGACATCTGTCCCCTCAGTTCTGC	213
O	y		67	TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer	86
D	b		214	TACGGCTCCAAGGCCCTACGTGAAGCACCCCACTGGTATCCCCGACTACTTCAAGCTGTCC	273
O	y		87	PheProGluGlyPheLysTrpGluArgValMetAsnPhenGluAspGlyGlyValValThr	106
D	b		274	TTCCCCGAGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAGCAGCGCGCTGTGTAAC	333
O	y		107	ValThrGlnAspSerSerLeuGlnAspGlyCySPheileTyrLysValLysPheileGly	126
D	b		334	GTCGCTCAGGACTCCTCCCTGCAGAGCGCGAATTCACTACAAGGTGAAGCTGCGCGGC	393
O	y		127	ValAsnPheProSerAspGlyProValMetGlnLysLysThrMerGlyTrpGluAlaSer	146
D	b		394	ACCAACTTCCCCCTCCGACGGCCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCTCC	453
O	y		147	ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyLulleHisLysAlaleuLys	166
D	b		454	TCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGATGAGGCTGAAG	513
O	y		167	LeuLysAspGlyGlyHIS TyrLeuValGluPheLysSerileTyrtMetalalysLysPro	186
D	b		514	CTGAAGGACGGCGGCCACTACAGCGCGGAGACCAAGACCACTTCAAGGCCAAGAAAGCCC	573
O	y		187	ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThxSerHisABngLu	206
D	b		574	GTGCAGTTGCCCGCGCCCTACATAAGCCGGCAGAAAGATGCATCATCCTCCCAATGAG	633
O	y		207	AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221	
D	b		634	GACTACACTATCGTGAATTGTACGAGCGCGCCGAGGCGGCCAC 678	

Search completed: January 12, 2006, 13:57:52  
Job time : 3773 secs



***THIS PAGE BLANK (USPTO)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 11:20:58 ; Search time 478 Seconds  
(without alignments)  
3137.146 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMKFKVMEG.....EDYTIQEYERTEGRHLEFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 83

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=rlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_10012006\_162503\_19106/app\_query.fasta\_1.391  
-DB=N\_Geneseq -QFMT=fastap -SUFFIX=p2nminscr.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=500 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=69 -ALIGN=500  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922@cgn\_1\_1\_1096@runat\_10012006\_162503\_19106 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_21: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*  
14: geneseqn2005s: \*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1214	100.0	678	4	Aah47654 Discosoma
2	1214	100.0	678	4	AAD13053 Discosoma
3	1214	100.0	678	4	AAD11142 Discosoma
4	1214	100.0	678	6	ABA95905 Yeast opt

c

5	1214	100.0	678	6	ABA95922	Abag95922 Yeast opt
6	1214	100.0	678	6	ABA95921	Abag95921 Yeast opt
7	1214	100.0	678	6	ABA95920	Abag95920 Yeast opt
8	1214	100.0	678	6	AAD46278	Aad46278 Discosoma
9	1214	100.0	678	6	AAD28207	Aad28207 Discosoma
10	1214	100.0	678	10	ADC24127	Adc24127 Discosoma
11	1214	100.0	678	10	ADF70404	Adf70404 Discosoma
12	1214	100.0	678	11	ADL46204	Adl46204 Discosoma
13	1214	100.0	678	11	ADN33978	Adn33978 Wild-type
14	1214	100.0	678	12	AD136420	Ad136420 Discosoma
15	1214	100.0	678	12	ADM97768	Adm97768 D sp red
16	1214	100.0	695	3	AAA48743	Aaa48743 Humanised
17	1214	100.0	859	6	AAL47952	Aal47952 Discosoma
18	1214	100.0	859	7	ADY51715	Ady51715 Discosoma
19	1214	100.0	859	8	AAD53432	Aad53432 Discosoma
20	1214	100.0	859	10	AAD61969	Aad61969 Discosoma
21	1214	100.0	859	14	ADX26533	Adx26533 Discosoma
22	1214	100.0	3311	9	ACA62995	Acac62995 Plasmid D
23	1210	99.7	681	10	ADC24134	Adc24134 Discosoma
24	1210	99.7	681	11	ADL46205	Adl46205 Human cod
25	1210	99.7	681	11	ADL46225	Adl46225 Human cod
26	1210	99.7	723	8	ABZ22476	Abz22476 Mammalian
27	1210	99.7	1638	10	ADL18131	Adl18131 RFP:PS (NI
28	1210	99.7	1647	10	ADL18155	Adl18155 RFP:PS (HI
29	1210	99.7	2721	13	ADS75466	Ad875466 Fibrohexa
30	1210	99.7	2772	13	ADS75468	Ad875468 Fibrohexa
31	1210	99.7	4692	6	AAL47954	Aal47954 Modified
32	1210	99.7	4692	10	ACC44640	Acc44640 Vector PD
33	1210	99.7	5436	4	AAD10003	Aad10003 Plasmid P
34	1210	99.7	6893	10	ADE24111	Ad24111 Proviral
35	1210	99.7	7910	4	AAD09979	Aad09979 pBit(dHSP
36	1210	99.7	9320	6	ABS56664	Ab556664 Plasmid P
37	1210	99.7	9658	12	AD104104	Ad104104 Vector PT
38	1210	99.7	9678	12	AD104103	Ad104103 Vector PT
39	1210	99.7	10263	12	AD104102	Ad104102 Vector PT
40	1209	99.6	681	6	ABA95906	Abag95906 Yeast opt
41	1209	99.6	681	6	ABA95907	Abag95907 Yeast opt
42	1207	99.4	678	6	AAD28208	Aad28208 Discosoma
43	1205	99.3	681	4	Aah47656	Aah47656 Anthozoan
44	1203	99.1	7508	13	ACR30814	Adr30814 Zebrafish
45	1197	98.6	6985	10	ACA55359	Acas5359 Transform
46	1196	98.5	666	6	ABL61142	Ab161142 Red fluor
47	1196	98.5	711	12	ADO78067	Ado78067 Corallimo
48	1196	98.5	711	12	ADO78069	Ado78069 Corallimo
49	1196	98.5	711	14	ADZ84222	Adz84222 Actinodis
50	1196	98.5	711	14	ADZ84224	Adz84224 Actinodis
51	1191	98.1	678	6	AAD46282	Aad46282 Discosoma
52	1191	98.1	678	6	AAD28209	Aad28209 Discosoma
53	1186	97.7	675	6	AAD46281	Aad46281 Discosoma
54	1186	97.7	747	14	ADV96699	Adv96699 Plant cel
55	1186	97.7	4200	14	ADW43589	Adw43589 UbCP-lox-
56	1186	97.7	4300	14	ADW43588	Adw43588 UbCP-lox-
57	1186	97.7	4488	13	ADQ91220	Adq91220 Fusion co
58	1186	97.7	6706	12	ADL35208	Adl35208 Plasmid P
59	1186	97.7	7495	13	ADR30815	Adr30815 Zebrafish
60	1186	97.7	7616	13	ADQ91222	Adq91222 Expressio
61	1186	97.7	7927	12	ADL35211	Adl35211 Plasmid P
62	1186	97.7	9731	11	ADM82791	Adm82791 DNA repai
63	1186	97.7	9782	11	ADM82792	Adm82792 DNA repai
64	1184	97.5	898	3	AAA48742	Aaa48742 Discosoma
65	1184	97.5	898	4	AAD03615	Aad03615 Discosoma
66	1160	95.6	678	10	ADC24128	Adc24128 Discosoma
67	1160	95.6	678	11	ADL46207	Adl46207 Discosoma
68	1160	95.6	704	11	ADN33980	Adn33980 Mutant -ty
69	1132	93.2	678	12	ADM13536	Adm13536 DNA encod
70	1121	92.3	681	10	ADC24130	Adc24130 Discosoma
71	1121	92.3	681	11	ADL46209	Adl46209 Discosoma
72	1119	92.2	678	3	AAA48748	Aaa48748 Humanised
73	1119	92.2	678	6	AAD46280	Aad46280 Discosoma
74	1111	91.5	678	3	AAA48749	Aaa48749 Humanised
75	1105	91.0	678	3	AAA48750	Aaa48750 Humanised
76	1085.5	89.4	876	3	AAA28958	Aaa28958 Discosoma
77	1085.5	89.4	876	4	AAD03618	Aad03618 Discosoma







XX Stack JH, Whitney M, Cubitt AB, Pollok BA;  
PI WPI; 2001-488890/53.  
XX Destabilizing proteins in living cells, by coupling a target protein to  
PT linear multimerized destabilization domain non-cleavable by -NH-ubiquitin  
PT protein endoproteases, comprising two copies of the domain.  
XX Disclosure; Page 110; 171pp; English.  
XX The present invention relates to a method for destabilizing a target  
CC protein in a cell. The method comprises a linker moiety which operatively  
CC couples a target protein (a reporter moiety) to a linear multimerised  
CC destabilising domain, which is non-cleavable by a -NH-ubiquitin protein  
CC endoproteases. The method is useful for detecting an activity such as  
CC protease, protein kinase or phosphoprotein phosphatase activity and is  
CC also useful for identifying modulators of these activities. The method is  
CC also useful for developing novel assays for a wide range of post-  
CC translational activities, such as proteolysis, phosphorylation,  
CC dephosphorylation, glycosylation, methylation, sulfation, prenylation,  
CC disulfide bond formation and ADP-ribosylation within cells. The  
CC recombinant DNA molecule of the invention is useful for creating  
CC transgenic animals useful as disease models and transgenic plants with  
CC improved disease resistance or other favourable traits. The present  
CC sequence is Discosoma sp. "red" anthozoa fluorescent protein, drFP483  
CC cDNA which is a natural fluorescent protein used as a reporter moiety in  
CC the exemplification of the invention

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-006-922A-12 (1-225) x AAD13053 (1-678)

OY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAAAGATGTTATCAAGAGGTCATGAGGTTTAAAGTTCCGATGGAAGCA 60  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGCAGAGTTTGAATAAGAGCGAAGAGAGAGCGGAGCCATACGAAGGC 120  
OY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGACCTTGGCCATTGCTGGGATATT 180  
OY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 TTGTCACCAACAATTCAGTATGGAAGCAAGGTATATGCAACCACTGCCGACATACCA 240  
OY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnGlu 100  
DB 241 GACTATAAAGAGCTGTCATTTCTGAAAGATTAAATGGGAAAGGGTCATGAACTTGAA 300  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCAAGATTCCAGTTGCAAGATGGCTGTTTCATCTAC 360  
OY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTGGCGTAACCTTCCCTCCGATGACCTGTTATGCAAAAGAAAGACA 420  
OY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCAGCACTGAGCGTTGTATCTCGTGAATGGCGTGTGAAGAAGAGAG 480

OY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 481 ATTCATTAAGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATCAAAAGTATT 540  
OY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAAAGCTGTGACGTACCAAGGGTACTACTATGTGACTCCAAACTGAT 600  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAAGAACCGAGGAGCG 660  
OY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCCTT 675  
RESULT 3  
AAD11142  
ID AAD11142 standard; DNA; 678 BP.  
XX  
AC AAD11142;  
XX  
DT 24-SEP-2001 (first entry)  
XX  
DE Discosoma species Anthozoa fluorescent protein, drFP583 encoding DNA.  
XX  
XX Transmembrane potential; biological membrane; fluorescent ion; detection;  
KW test chemical screening; Anthozoa fluorescent protein; FP;  
KM transgenic organism; drFP583 protein; ds.  
XX  
OS Discosoma sp.  
XX  
PN WO200142211-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 12-DEC-2000; 2000WO-US033739.  
XX  
PR 13-DEC-1999; 99US-00459956.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Gonzalez JE;  
XX  
DR WPI; 2001-457276/49.  
XX  
XX Determining electrical potential across a membrane in biological systems,  
PT comprises introducing two reagents, exposing the membrane to light and  
PT measuring the energy transfer.  
XX  
PS Disclosure; Page 150; 154pp; English.  
XX  
XX The patent discloses optical methods and compositions for determining  
CC transmembrane potential across biological membranes in living cells. The  
CC method of determining the electrical potential across a membrane in a  
CC biological system comprises introduction of two reagents, a first reagent  
CC comprising a hydrophobic fluorescent ion capable of redistributing from a  
CC first face of the membrane to a second face of the membrane in response  
CC to membrane potential change and a second reagent that label the first  
CC face or the second face of the membrane, which comprises a chromophore  
CC capable of undergoing energy transfer by either donating or accepting  
CC excited state energy to the fluorescent ion. The membrane is then exposed  
CC to excitation light and the energy transfer between the reagents is  
CC measured and related to the membrane potential. The method is useful for  
CC detecting changes in membrane potential in subcellular organelle  
CC membranes in biological systems. The method is used for screening of test  
CC chemicals for activity to modulate the activity of target ion channel.  
CC The invention also provides a transgenic organism comprising a first  
CC reagent that comprises a charged hydrophobic fluorescent molecule and a  
CC second reagent comprising a bioluminescent or naturally fluorescent  
CC protein. The present sequence is Discosoma species (red) DNA encoding an  
CC Anthozoa fluorescent protein (FP), drFP583



SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-006-922A-12 (1-225) x AAD11142 (1-678)

QY 1 MetArgSerSerLySAsnValIlleYsGluPheMetArgPheLySValArgMetGluGly 20

DB 1 ATGAGGTTCTTCCAAGATGTTATCAAGAGTTCATGAGGTTTAAGGTTCCGATGGAAGCA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

DB 61 ACGGTCAATGGGACGAGTTTGAAATGAAGCGAAGAGAGGGAGGCCATACGAAGGC 120

QY 41 HisAsnThrValLySLeuLySValThrLySGLYGLYProLeuProPheAlaTyrAspIle 60

DB 121 CACAATACCGTAAGAAGCTTAAGGTAACCAAGGGGGACCTTGGCCATTGCTTGGAATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80

DB 181 TTGTCAACCAATTTCAAGTATGAGCAAGGTATATGTCAAGCACCTGCCGACATACCA 240

QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPheGlu 100

DB 241 GACTATATAAAGCTGTCAATTTCTGAAGCATTTAAATGGGAAGGCTCATGAACCTTGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB 301 GACGTGGCGTCTTACTGTACGTAACCCAGGATTCAGTTGCAGGATGGCTGTTTCATCTAC 360

QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140

DB 361 AAGGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAAGACA 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLYlu 160

DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGCTTGTATCTCGTGATGGCGTGTGAAGAGAGAG 480

QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180

DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

DB 541 TACATGGCAAGAAGCCTGTGTCAGCTACCAAGGTTACTACTATGTGACTCCAAACTGGAT 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 601 ATTAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAAGCCGAGGGAGCGC 660

QY 221 HisHisLeuPheLeu 225

DB 661 CACCATCTGTCTCCTT 675

RESULT 4

ABA95905 standard; cDNA; 678 BP.

XX ABA95905;

XX 29-MAY-2002 (first entry)

DE Yeast optimised RFP encoding cDNA SEQ ID NO 1.

XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;

KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;

KX Escherichia coli; green fluorescent protein; biotechnology; gene; ss.

XX Anthozoa.

OS

XX Key

FH CDS

FT 1..678

FT /\*tag= a

FT /product= "Yeast optimised Red Fluorescent Protein"

XX DE20001395-U1.

PN

XX 15-MAR-2001.

PD

XX 27-JAN-2000; 2000DE-02001395.

PF

XX 27-JAN-2000; 2000DE-02001395.

PR

XX (GPCB-) GPC BIOTECH AG.

PA

XX WPI; 2002-228394/29.

XX P-PSDB; ABB08834.

DR

XX New DNA encoding red fluorescent protein, useful as marker in

PT biotechnology, has sequence optimized for expression in eukaryotes,

PT especially yeast or plants.

PT

XX Claim 1; Fig 1; 19pp; German.

PS

XX The invention relates to DNA (I) containing either sequence ABA95905 or

CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein

CC (YRFP). (I) are used to express red fluorescent protein (RFP) in

CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and

CC plants, especially dicotyledonous plants including Nicotiana tabacum or

CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,

CC especially Escherichia coli. RFP is useful in the same way as green

CC fluorescent protein but is more generally applicable in modern

CC biotechnology. (I) are optimised for expression in yeast and so generate

CC RFP at higher levels with stronger fluorescence and thus lowers the

CC detection limit and gives a better signal-to-noise ratio

XX

SQ Sequence 678 BP; 198 A; 147 C; 159 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95905 (1-678)

QY 1 MetArgSerSerLySAsnValIlleYsGluPheMetArgPheLySValArgMetGluGly 20

DB 1 ATGAGATCTTCTAAGACGTGATCAAGGAATTCATGAGATTCAAGTTAGAATGAAGGT 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

DB 61 ACTGTTAACGGTCACGAATTCGAAATCGAAGGTGAAGGTGAAGTAGACCATACGAAGGT 120

QY 41 HisAsnThrValLySLeuLySValThrLySGLYGLYProLeuProPheAlaTyrAspIle 60

DB 121 CACAACACTGTCAAGTTGAAGTTACTTAAGGGTGGTGCATTGGCCATTGCTTGGGACATC 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80

DB 181 TTGTCTCCACAATTCCAATACGTTCTTAAGGTCTACGTCAAGCACCCAGCTGACATTCCA 240

QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpgLuArgValMetAsnPheGlu 100

DB 241 GACTACAAGAAGTGTCTTCCCAAGAAGTTTCAAGTGGGAAGAAGTCATGAACCTTGGAA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120



Db 301 GACGGTGGTGTGTTACTGTACTCAAGACTCCTCTTGCAAGACGGTGTTCATCTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGTCAAGTTCATCGGTGTCAACTCCCATCTGACGGTCCAGTCAATGCAAAAGAACT 420  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 421 ATGGGTGGGAAGCTTCTACCGAAGCTTTGTAACCAAGAGACGGTGTGAAAGGTGAA 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGGCTTGAAGTTGAAGACGGTGTGCTACTACTGTGCAATTCAAGTCTATC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCTAAGAAAGCCAGTCCAAATTGCCAGGTTACTACTACGTTGACTCTAAGTTGAC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACCCTCTCACAAAGAACTACACTATCGTGAACAATACGAACGTTACTGAAGGTAGA 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTTGTTCTTG 675

RESULT 5  
ABA95922  
ID ABA95922 standard; DNA; 678 BP.

XX ABA95922;  
AC  
XX 29-MAY-2002 (first entry)

XX Yeast optimised RFP related DNA SEQ ID NO 19.

DE Yeast optimised RFP related DNA SEQ ID NO 19.  
XX  
XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KM Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KM Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
OS Anthozoa.

XX Key Location/Qualifiers  
FH 1..678  
FT /\*tag= a  
FT /product= "Yeast optimised Red Fluorescent Protein"

FT DE20001395-U1.

XX 15-MAR-2001.

XX 27-JAN-2000; 2000DE-02001395.

XX 27-JAN-2000; 2000DE-02001395.

XX (GPCB-) GPC BIOTECH AG.

XX WPI; 2002-228394/29.

XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.

XX Disclosure; Page 14-15; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern

CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
CC useful to the invention

XX  
SQ Sequence 678 BP; 202 A; 118 C; 159 G; 199 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95922 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGATCTTCTAAGAAGCTTATTAGAATTCATGAGATTCAAGGTTAGAATGGAAGGT 60

Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACTGTTAACGGTCACGAATTCGAATTTGAAGTGAAGGTGAAGGTAGACCATACGAAGGT 120

Qy 41 HisAsnThrValLysLeuLysValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACTGTTAAGTTGAAGGTTACTAAGGTTGTCATTGCCATTGCTGGGACATT 180

Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGCTCCACAATTCCAATACGGTTCTTAAGTTTACGTTAAGCACCCAGTGAACATTCCA 240

Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAGTTGTTCTTCCAGAAGTTTCAAGTGGAAAGAGTTATGAACCTTCGAA 300

Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGTGTTGTTACTGTTACTCAAGACTCTTCTTGCAAGACGGTGTGTTCAATTAC 360

Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTTAAGTTCAITGGTGTTAACCTTCCATCTGACGGTCCAGTTATGCAAAAGAACT 420

Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 421 ATGGGTGGGAAGCTTCTACTGAAGAATGTACCAAGAGACGGTGTGTTGAAGGGTGAA 480

Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCACAAGGCTTGAAGTTGAAGACGGTGTCACTACTGTTGAATTCAAGTCTTATT 540

Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCTAAGAACCACTTCAATTGCCAGGTTACTACTACGTTGACTCTAAGTTGAC 600

Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATTACTTCTCACAAAGAACTACACTATTGTTGAACAATACGAAGAAGACTGAAGTAGA 660

Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCACTTGTTCTTG 675

RESULT 6  
ABA95921  
ID ABA95921 standard; DNA; 678 BP.

XX ABA95921;  
AC  
XX 29-MAY-2002 (first entry)



XX Yeast optimised RFP related DNA SEQ ID NO 18.  
DE Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology; gene; ss.  
XX Anthozoa.  
OS Anthozoa.  
XX Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "yeast optimised Red Fluorescent Protein"  
XX DE20001395-UI.  
XX 15-MAR-2001.  
XX 27-JAN-2000; 2000DE-02001395.  
XX 27-JAN-2000; 2000DE-02001395.  
XX 27-JAN-2000; 2000DE-02001395.  
XX (GPCB-) GPC BIOTECH AG.  
XX WPI; 2002-228394/29.  
XX P-PSDB; ABB08834.  
XX  
XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
XX Disclosure; Page 14; 19pp; German.  
XX  
XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or  
CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (I) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of a polynucleotide encoding the yeast optimised RFP,  
CC useful to the invention  
XX  
XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABA95921 (1-678)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGCTTCCAAAGATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCCATGGAAGA 60  
QY 21 ThrValasnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrglyGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGAGAGAGGGGACCATACGAAGGC 120  
QY 41 HisAsnThrVallySleuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGGCTTGGAATATT 180  
QY 61 LeuSerProGlnPheGlnTyrglySerLySValTyrrVallyHisProAlaAspIlePro 80

DB 181 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTCGCATATACCA 240  
QY 81 AspTyrLySlySleuSerPheProGluGlyPheLySTrpGluArgValMetAsnPhelGly 100  
DB 241 GACTATTAATAAAGCTGTCAATTCTGGAAGATTAAATGGGAAGGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCGTACTGTATACCCAGGATTCAGATTGCGAGATGGCTGTTTCATCTAC 360  
QY 121 LySVallySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThr 140  
DB 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGACTGTTATGCCAAAAGAGACA 420  
QY 141 MetGlyTrrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAATGGCGTGTGAAGGAGAG 480  
QY 161 IleHisLySAlaLeuLySleuLySAspGlyGlyHisSTyrLeuValGluPheLySserIle 180  
DB 481 ATTCAATAAGGCTCTGAAGCTGAAGACGGTGTTCATTACCTAAGTTCAAAAGTATT 540  
QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySleuAsp 200  
DB 541 TACATGGCAAGAAGCCTGTGACGCTACCAAGGTAAGTGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrglyuArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACGAAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCTT 675

RESULT 7  
ABA95920  
ID ABA95920 standard; RNA; 678 BP.  
XX ABA95920;  
XX 29-MAY-2002 (first entry)  
XX  
XX Yeast optimised RFP encoding RNA SEQ ID NO 16.  
DE Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;  
KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;  
KW Escherichia coli; green fluorescent protein; biotechnology; ss.  
XX Anthozoa.  
OS Anthozoa.  
XX DE20001395-UI.  
XX 15-MAR-2001.  
XX 27-JAN-2000; 2000DE-02001395.  
XX 27-JAN-2000; 2000DE-02001395.  
XX 27-JAN-2000; 2000DE-02001395.  
XX (GPCB-) GPC BIOTECH AG.  
XX WPI; 2002-228394/29.  
XX  
XX New DNA encoding red fluorescent protein, useful as marker in  
PT biotechnology, has sequence optimized for expression in eukaryotes,  
PT especially yeast or plants.  
XX  
XX Disclosure; Page 13; 19pp; German.  
XX  
XX The invention relates to DNA (I) containing either sequence ABA95905 or  
CC sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein  
CC (yRFP). (I) are used to express red fluorescent protein (RFP) in  
CC eukaryotes, especially yeast, especially Saccharomyces cerevisiae and  
CC plants, especially dicotyledonous plants including Nicotiana tabacum or



CC Arabidopsis thaliana and also in prokaryotes, especially bacteria,  
CC especially Escherichia coli. RFP is useful in the same way as green  
CC fluorescent protein but is more generally applicable in modern  
CC biotechnology. (1) are optimised for expression in yeast and so generate  
CC RFP at higher levels with stronger fluorescence and thus lowers the  
CC detection limit and gives a better signal-to-noise ratio. The present  
CC sequence is that of an RNA sequence corresponding to the yeast optimised  
CC RFP encoding cDNA given as SEQ ID NO 1 (ABA95905)

XX SQ Sequence 678 BP; 198 A; 147 C; 159 G; 0 T; 174 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x ABA95920 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPhelyValArgMetGluGly	20
DB	1	AUGAGAUCCUUCUAAAGAACGUCUACAAGGAUUCAGAUAUCCAAGGUAAGAAGGU	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACUGUUAACGGUCACGAUUCGAAAUCCGAAGGUGAAGGUAAGGUAACCAUACGAAGGU	120
QY	41	HisAsnThrValLyseuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACACACACUGUACAAGUUGAAGUUAACUAAAGGUGUCCAUUGCCAUUGCGUUGGAGACAU	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro	80
DB	181	UUGUCUCCACAUAUCCAUACGGUUCUAAAGUUCAGUCAAGCACCCAGCUGACAUUCCA	240
QY	81	AspTyrLyLySLeuSerPheProGluGlyPheLySerTrpGluArgValMetAsnPhelGlu	100
DB	241	GACUACAAAGAGUUGUCCUCCCAAGAGGUUUCAGUGCGAAGAAGUACAUGACUUCGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	301	GACGGUGUGUUGUUAUCGUUACUCAAAGACUCCUCCUCCGACAAGACGGUUGUUAUCUAC	360
QY	121	LySValLySPheIleGlyValAsnPhProSerAspGlyProValMetGlnLySlyThr	140
DB	361	AAGGUCAAAGUUAUCGGUGUCAAUCCCAUUCGACGGUCCAGUACAUGCAAAAAGAGACU	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160
DB	421	AUGGGUUGGGAAGCUUUAACCGAACGUUUGUACCAAGAGACGGUGUCUUGAAGGGUGAA	480
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle	180
DB	481	AUCCACAAGGCCUUGAAGUUGAAGACGGUGGUCAUACUUGGUGCGAAUUCAGAUCUAC	540
QY	181	TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	541	UACAUGGCUAAGAACCAAGUCCAUUUGCCAGGUUACUACGUUGACUCUAAAGUUGGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	601	AUCAACUUCUACAACGAAGACUACUUAUCGUCCGAACAUAACGAACGUACUGAAGGUA	660
QY	221	HisHisLeuPheLeu 225	
DB	661	CACCAUCUGUUCUUG 675	

RESULT 8  
AAD46278  
ID AAD46278 standard; DNA; 678 BP.  
XX

AC	AAD46278;
XX	
DT	27-DEC-2002 (first entry)
XX	
DE	Discosoma sp. drFP583 (NFP-6) wild-type protein encoding DNA.
XX	
KW	Fluorescent protein; chromoprotein; protease cleavage assay; filter;
KW	fluorescence activated cell sorting application; fluorescent timer;
KW	biosensor; fluorescence resonance energy transfer application; FRET;
KW	colouring agent; recombinant DNA application; analyte detection assay;
KW	sunscreen; second messenger detector; drFP583 protein; NFP-6; gene; ds.
XX	
OS	Discosoma sp.
XX	
FH	Key Location/Qualifiers
FT	1..678
FT	/*tag= a
FT	/product= "drFP583 wild-type protein"
XX	
PN	WO200268459-A2.
XX	
PD	06-SEP-2002.
XX	
PF	20-FEB-2002; 2002WO-US005749.
XX	
PR	21-FEB-2001; 2001US-0270983P.
PR	04-DEC-2001; 2001US-00006922.
XX	
PA	(CLON-) CLONTECH LAB INC.
XX	
PI	Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;
XX	
DR	WPI: 2002-691654/74.
DR	P-PSDB; AAE28833.
XX	
PT	New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant
PT	of an aggregating Cnidarian chromo- or fluorescent protein or mutant for
PT	analyte detection assays or fluorescence activated cell sorting
PT	applications.
XX	
PS	Disclosure; Page 70; 80pp; English.
XX	
CC	The invention relates to nucleic acid molecules encoding non-aggregating
CC	chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are
CC	useful in analyte detection assays, as colouring agents, as markers in
CC	recombinant DNA applications, as sunscreens or filters, in fluorescence
CC	resonance energy transfer (FRET) applications, as biosensors in
CC	prokaryotic and eukaryotic cells, in screening assays, as second
CC	messenger detectors, in fluorescence activated cell sorting applications,
CC	in protease cleavage assays or as fluorescent timers. The present
CC	sequence is a DNA encoding Discosoma sp. drFP583 (NFP-6) wild-type
CC	protein of the invention
XX	
SQ	Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AAD46278 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPhelyValArgMetGluGly	20
DB	1	ATGCGCTCCTCCAGAAGCGTCATCAAGAGTTGCGCTTCAAGGTGCGCATGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGGC	120



Qy	41	HisAsnThrValIysLeuLysValThrLysGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACAACACCGTGAAGCTGAAGGTGAACCAAGGGCGGCCCTTCGCTTGCGACATC	180
Qy	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCAGTAGACGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCC	240
Qy	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACCTTCGAG	300
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGGCGTGTGACCGTGACCCAGACTCTCTCCCTGCAGAGCGGCTGCTTCATCTAC	360
Qy	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTGAAGTTCAATCGGCGTGAACCTTCCTCCGACGCGCCCGTGAATGAGAAGAACCC	420
Qy	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGCGTGTGAAGGGCGAGC	480
Qy	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTCTGTGAGTTCAAGTCCATC	540
Qy	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCAAAGACCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAAGCTGAC	600
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACTCCCAACAAGGACTACACCACTGTGGAGCAGTACGAGCGCACCGAGGGCGGC	660
Qy	221	HisHisLeuPheLeu	225
Db	661	CACCACCTGTTCTCTG	675
RESULT 9			
AAD28207			
ID	AAD28207 standard; DNA; 678 BP.		
AC	AAD28207;		
XX			
DT	22-APR-2002 (first entry)		
XX			
DE	Discosoma sp. humanised wild-type Anthozoa protein drFP583 DNA.		
XX			
KW	Fluorescent timer protein; protein movement; translocation; trafficking;		
KW	promoter activity; gene expression; transgenic plant; gene modification;		
KW	protein age; anthozoa protein; drFP583; ds.		
OS	Discosoma sp.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..678	
FT	/*tag= a		
FT	/product= "Humanised wild-type Anthozoa protein drFP583"		
XX			
PN	WO200196373-A2.		
XX			
PD	20-DEC-2001.		
XX			
PF	13-JUN-2001; 2001WO-US019097.		
XX			
PR	14-JUN-2000; 2000US-0211607P.		
XX			
PA	(CLON-) CLONTECH LAB INC.		
XX			
PI	Fradekov AF, Tersikh A;		
XX			

DR

WPI; 2002-154595/20.

DR

P-PSDB; AAE17540.

XX

PT

New fluorescent timer proteins comprising an emission spectrum that

PT

changes over time from a first wavelength to a second wavelength, useful

PT

for monitoring intracellular protein movement, translocation, trafficking

PT

or stability.

XX

PS

Example 1; Fig 1; 89pp; English.

XX

CC

The invention relates to a fluorescent timer protein having an emission

CC

spectrum that changes over time after synthesis from a first wavelength

CC

to a second wavelength. The fluorescent timer proteins are useful in

CC

monitoring the activity of a promoter, determining the age of a protein,

CC

identifying an agent that modulates the activity of a promoter and in

CC

enriching a population of cells comprising a fluorescent timer protein.

CC

The fluorescent timer proteins are also useful for assessing gene

CC

expression during development of a multicellular organism or during

CC

cellular differentiation, in response to a drug or other inducer of

CC

promoter activity, as a reporter to serve as a read-out of promoter

CC

activity, monitoring intracellular protein movement or translocation,

CC

protein trafficking, or protein stability, to investigate temporal

CC

aspects of the activity of a regulatory element, for determining cell

CC

fate during development and organ remodelling, in spatial and temporal

CC

visualisation of newly synthesised proteins and accumulated proteins, and

CC

in distinguishing between newly formed and pre-existing structures, e.g.

CC

membrane junctions and extracellular matrix components. The fluorescent

CC

timer proteins may further be used to investigations where photobleaching

CC

techniques are employed, as detectable labels, as selectable markers, as

CC

biosensors in prokaryotic and eukaryotic cells, in protease cleavage

CC

assays, and as second messenger detectors. The nucleic acids can be used

CC

to generate transgenic, non-human plants or animals or site-specific gene

CC

modifications in cell lines. The present sequence is a DNA encoding

CC

Discosoma sp. humanised wild-type Anthozoa protein drFP583 used for

CC

generating fluorescent proteins

XX

SQ

Sequence 678 BP; 146 A; 223 C; 203 G; 106 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

2.3e-140

Length: 678

Score:

1214.00

Matches: 225

Percent Similarity:

100.00%

Conservative: 0

Best Local Similarity:

100.00%

Mismatches: 0

Query Match:

100.00%

Indels: 0

DB:

6

Gaps: 0

US-10-006-922A-12 (1-225) x AAD28207 (1-678)

QY

1 Met

ArgSerSerIySAsnValIleIySGluPheMetArgPheIySValArgMetGluGly

20

Db

1 ATG

CGCTCTCCAAGAACGTATCAAGAGTTCATCGCTTCAAGGTGCGCATGAGGGC

60

QY

21 Thr

ValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly

40

Db

61 ACC

GTGAACGGCCACGAGTTCAGATCGAGGCGGAGGGCGGCCCTTACGAGGGC

120

QY

41 His

A

antThrValIlySLeuIySValThrIySGlyProLeuProPheAlaTrpAspIle

60

Db

121 CACA

ACACCGGTGAAGCTGAAGGTGAACCAAGGCGGCCCTTCGCTTGAGACATC

180

QY

61 Leu

SerProGlnPheGlnTyrGlySerIySValTyrValIySHisProAlaAspIlePro

80

Db

181 CTGT

CCCCCAGTTCAGTAGCGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC

240

QY

81 Asp

TyrlIySLeuSerPheProGluGlyPheIySTrpGluArgValMetAsnPheGlu

100

Db

241 GACT

ACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG

300

QY

101 Asp

GlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

Db

301 GACG

CGCGCGGTGTGACCGTGACCCAGACTCTCCCTGCAGAGACGGCTGCTTCATCTAC

360

QY

121 Lys

ValIySLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr

140



|||||  
DB 361 AAGTGAAGTTCATCGCGGTGAACCTCCCTCCGACGGCCCGTGATGCAGAAAGACC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCCACCGAGCGCTGTACCCCCCGACGGCGTGTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCCACAAAGGCCCTGAAGCTGAAGAGCGCGGCCACTACTGTGTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAGAGCCCGTGACAGTGGCCGCTACTACTACGTGACTCCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAAGACGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACCTGTCTCTG 675  
RESULT 10  
ADC24127  
ID ADC24127 standard; DNA; 678 BP.  
XX ADC24127;  
AC  
XX  
XX 18-DEC-2003 (first entry)  
DT  
XX  
DE Discosoma wild-type red fluorescent protein DNA #1.  
XX  
XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KM fluorescent protein variant; transcription induction detection;  
KM fluorescence energy resonance transfer; FRET; protein kinase;  
KM protein phosphatase; ion indicator; ds.  
XX  
OS Discosoma.  
XX  
XX US2003059835-A1.  
PN  
XX  
XX 27-MAR-2003.  
PD  
XX  
XX 10-APR-2002; 2002US-00121258.  
PF  
XX  
XX 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
PR  
XX  
XX (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
PA  
XX  
XX Tsien RY, Campbell RE;  
PI  
XX  
XX WPI; 2003-743764/70.  
DR P-PSDB; ADC24126.  
XX  
XX  
PT Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
PT  
XX  
XX Example 1; SEQ ID NO 2; 67bp; English.  
PS  
XX  
XX The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant

CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca<sup>2+</sup>, Zn<sup>2+</sup>, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma wild-  
CC type red fluorescent protein.  
XX  
SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-10-006-922A-12 (1-225) x ADC24127 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGGTCTTCCAGAGATGTATATCAAGAGATTCAAGGTTTAAGGTTGCGATGGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyArgProGlyGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTGAATAGAAAGCGCAAGAGAGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTACCAACAATTTCAATGGAAGCAAGGTATATGCAAGCACCCTGCCAGATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnDheGlu 100  
DB 241 GACTATAAAAGCTGTGCTTTCTCGAAGATTTAATGGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCTTACTGTAAACCAAGATTCCAGTTTCCAGGATGGCTGTTCACTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGACAGTGAAGCGTTGTATCCTGTGATGGCGTGTGAAGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTACTAATTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAGAGCCGTGACAGCTACCAAGGTAATGTTGACTCCAAGCTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220



DB
601
ATACCAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCGAGGAGCGC
660

QY
221
HISH1eLeuPheLeu
225
661
CACCATCTGTTCCTT
675

RESULT 11
ADP70404
ID
ADF70404
standard; cDNA; 678 BP.

XX
AC
ADF70404;
XX
DT
12-FEB-2004
(first entry)
XX

DE
Discosoma wild-type GFP variant cDNA SeqID27.

XX
KW
ligand; orphan receptor protein; fusion protein; fluorescent protein; cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP; GFPuv; Enhanced GFP; EGFP; gene; ss.

XX
OS
Discosoma sp.

XX
PN
WO2003071272-A1.

XX
PD
28-AUG-2003.

XX
PF
21-FEB-2003; 2003WO-JP001901.

XX
PR
22-FEB-2002; 2002JP-00045728.

XX
PR
23-JUL-2002; 2002JP-00213949.

XX
PR
11-OCT-2002; 2002JP-00298237.

XX
PA
(TAKE ) TAKEDA CHEM IND LTD.

XX
PI
Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;

XX
DR
WPI; 2003-697654/66.

XX
DR
P-PSDB; ADF70403.

XX
PT
Transformation of cells with a fusion protein of an orphan receptor protein with a fluorescent protein useful for identification of ligands to the orphan receptor.

XX
PS
Disclosure; SEQ ID NO 27; 594bp; Japanese.

XX
CC
This invention relates to a novel method of identifying ligands to an orphan receptor protein which comprises transforming cells with DNA encoding a fusion protein of the orphan receptor with a fluorescent protein, so that the fusion protein is expressed in the cells (or cell membranes isolated from them) and contacting the cells with the potential ligand to be tested. A suitable fluorescent protein for incorporation in the fusion protein is green fluorescent protein (GFP), for example GFP-1, wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the identification of ligands binding to an orphan receptor protein.

SQ
Sequence 678 BP; 204 A; 129 C; 179 G; 166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
2.3e-140
Length:
678
Score:
1214.00
Matches:
225
Percent Similarity:
100.00%
Conservative:
0
Best Local Similarity:
100.00%
Mismatches:
0
Query Match:
100.00%
Indels:
0
DB:
10
Gaps:
0

US-10-006-922A-12 (1-225) x ADF70404 (1-678)

QY
1
MetArgSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly
20
1
ATGAGGTCTTCCAGAATGTATCAAGAGATGCATGAGGTTTAAGGTTGCGATGGAAGA
60
21
ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly
40

DB
61
ACGGTCAATGGGCACGAGTGTGAATAAGAGCGCAAGAGAGGGGAGCCATACGAAGGC
120

QY
41
HISAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTPAspIle
60
121
CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGGCAATTTGCTGGGATATT
180

QY
61
LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro
80
181
TTGTACACCACAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCTCGCACATACCA
240

QY
81
AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu
100
241
GACTATATAAAAGCTGTCATTTCCCTGAAGATTAAATGGGAAAGGTCATGAACCTTGAA
300

QY
101
AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr
120
301
GACGGTGGCGTCGTTACTGTATACCCAGGATTCAGTTTGCAAGATGGCTTTTCATCTAC
360

QY
121
LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr
140
361
AAGGTCAAGTTCAATTGGCGTGAACCTTCTCCGATGGACCTGTTATGCCAAAAGAGACA
420

QY
141
MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu
160
421
ATGGGCTGGGAAGCCAGCACGTGAGCGTTGTATCCTCGTGAATGGCGTGTGAAAGAGAG
480

QY
161
IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle
180
481
ATTCAATAAGGCTCTGAAGCTGAAGAAGCGTGTCACTTACCTAGTTGAATTCAAAAGTATT
540

QY
181
TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp
200
541
TACATGGCAAGAAGACCTGTGACGACTACCAAGGTACTACTATGTGACTCCAAACTGGAT
600

QY
201
IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg
220
601
ATACCAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAAGAACCGAGGAGCGC
660

QY
221
HISH1eLeuPheLeu
225
661
CACCATCTGTTCCTT
675

RESULT 12
ADL46204
ID
ADL46204
standard; DNA; 678 BP.

XX
AC
ADL46204;

XX
DT
20-MAY-2004
(first entry)
XX

DE
Discosoma red fluorescent protein (DsRed) coding sequence.

XX
KW
ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength; oligomerization; tetramerization; immunoassay; hybridization assay.

XX
OS
Discosoma sp.

XX
FH
Key
Location/Qualifiers
FT
CDS
1..678
/\*tag= a
FT
/product= "DsRed protein"

XX
PN
WO2003086446-A1.

XX
PD
23-OCT-2003.

XX
PF
09-APR-2003; 2003WO-US010879.

XX
PR
10-APR-2002; 2002US-00121258.

XX
PR
29-JUL-2002; 2002US-00209208.

XX
PA
(REGC ) UNIV CALIFORNIA.



XX Tsielen RY, Campbell RE, Baird GS;  
XX WPI; 2003-845265/78.  
DR P-PSDB; ADL46203.  
XX  
XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
XX Disclosure; SEQ ID NO 2; 166pp; English.  
XX  
XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC coding sequence.  
XX  
XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 2.3e-140 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46204 (1-678)

OY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 1 ATGAGGTCCTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAAGTTCGCATGGAAGCA 60  
OY 21 ThrValAsnGluHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAAATGAAAGGCGAAGGAGAGGGGACATACGAAGGC 120  
OY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrPheIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGACCTTGCCATTGGCTTGGAATATT 180  
OY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValIlyHisProAlaAspIlePro 80  
DB 181 TTGTCAACCACAATTTCAGTATGGAAGCAAGATATGTCAAGCACCTGCGACATACCA 240  
OY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpgLuarGValMetAsnPhedu 100  
DB 241 GACTATAAAAAAGCTGTCAATTCCTGAAGGATTAAATGGGAAGGGTCATGAACCTTGA 300  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCGTACTGTAAACCAAGGATTCAGTTTGAGATGGCTGTTTCATCTAC 360  
OY 121 LysValIlyAspPheIleGlyValAsnPhProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGACCTGTATGCAAAAGAAAGCA 420  
OY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyIu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAAGCGGTGTGAAGAGAGAG 480  
OY 161 IleHisLeuAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180

DB 481 ATTCAATAGGCTCTGAAGCTGAAGACGGTGTCTATTACCTAGTTGAATTCAAAAGTATT 540  
OY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAAAGCCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAGTGGAT 600  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAAGCCACAACGAAGACTATACATCTTGAAGCAGTATGAAGAACCAGGAGCGC 660  
OY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCTT 675  
RESULT 13  
ADN33978  
ID ADN33978 standard; DNA; 678 BP.  
XX  
XX ADN33978;  
AC  
XX 01-JUL-2004 (first entry)  
DT  
XX  
XX Wild-type DsRED encoding sequence.  
DE  
XX Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRED;  
KW ds.  
XX  
XX Discosoma sp.  
OS  
XX  
XX WO2003054158-A2.  
PN  
XX 03-JUL-2003.  
PD  
XX 18-DEC-2002; 2002WO-US040539.  
PF  
XX 19-DEC-2001; 2001US-0341723P.  
PR  
XX  
XX (UYCH-) UNIV CHICAGO.  
PA  
XX  
XX Bevis B, Glick B;  
PI  
XX WPI; 2003-569236/53.  
XX  
XX P-PSDB; ADN33979.  
DR  
XX  
XX Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
PT mutant of a Cnidarian chromo- or fluorescent protein or its mutant,  
PT useful for applications involving chromo- or fluorescent proteins.  
PT  
XX  
XX Claim 8; SEQ ID NO 1; 65pp; English.  
PS  
XX  
XX The present invention relates to nucleic acid that encodes a rapidly  
CC maturing chromo or fluorescent mutant of a Cnidarian chromo- or  
CC fluorescent protein or its mutant. The protein is useful in applications  
CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
CC useful for producing a chromo and/or fluorescent protein which involves  
CC growing the cell, whereby the protein is expressed, and isolating the  
CC protein substantially free of other proteins. The protein is useful in  
CC applications involving chromo- or fluorescent protein and is useful as  
CC PCR primers, hybridization probes, etc. The expression cassettes are  
CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for  
CC biological analytes of interest and as selectable markers in recombinant  
CC DNA applications, e.g. the production of transgenic cells and organisms.  
CC The fluorescent proteins find use in a variety of different applications,  
CC e.g. in fluorescence resonance energy transfer (FRET) applications, as  
CC biosensors in prokaryotic and eukaryotic cells, in applications involving  
CC the automated screening of arrays of cells expressing fluorescent  
CC reporting groups by using microscopic imaging and electronic analysis, as  
CC second messenger detectors, and in fluorescence activated cell sorting  
CC applications and as in vivo marker in animals. The fluorescent proteins



CC also find use in protease cleavage assays. The proteins can also be used  
CC is assays to determine the phospholipid composition in biological  
CC membranes and as a fluorescent timer. The present sequence represents the  
CC wild-type DsRED encoding sequence.

XX Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922A-12 (1-225) x ADN33978 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
DB	1	ATGAGGCTTCCAAAGATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCCGATGGAAGCA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACGGTCAATGGGACGAGTTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC	120
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACAATACCGTAAAGCTTAAAGTAACCAAGGGGGACCTTGGCATTTGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHisProAlaAspIlePro	80
DB	181	TTGTCAACCAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCCCTGCCACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	241	GACTATAAAAGCTGTCTATTCTCGAAGGATTTAAATGGAAAGGTCATGAACCTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	301	GACGGTGGCGTCGTACTGTAAACCAAGATTCAGTTTGACAGATGGCTGTTTCATCTAC	360
QY	121	LySValLySPhelIeGlyValAsnPheProSerAspGlyProValMetGlnLySThr	140
DB	361	AAGGTCAAGTTCATGGCGTGAACTTCCCTCCGATGACCTGTTATGCAAAAGAGACA	420
QY	141	MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160
DB	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGTGTGAAGAGAG	480
QY	161	IleHisLySValLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySSerIle	180
DB	481	ATTCAATAAGGCTCTGAAGCTGAAAGACGGTGTCTATTACCTAGTTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	541	TACATGGCAAGACCTGTGTGCAAGTACCAAGGTACTACTATGTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	601	ATAACAAGCCACAAAGAACTATACAATCGTTGAGCAGTATGAAGAAGCAAGGAGCGC	660
QY	221	HisHisLeuPheLeu	225
DB	661	CACCATCTGTCTCTT	675

RESULT 14  
AD136420  
ID AD136420 standard; DNA; 678 BP.

XX AC AD136420;  
XX DT 22-APR-2004 (first entry)  
XX

DE Discosoma sp. red fluorescent protein (RED) DNA.  
XX  
XX Fluorobody; binding ligand; green fluorescent protein; GFP;  
KW target detection; gene; red fluorescent protein; db; RED.  
XX

OS Discosoma sp.

Key	Location/Qualifiers
FT CDS	1..678
FT	/*tag= a
FT	/product= "Discosoma sp. red fluorescent protein (RED) "

US2003203355-A1.

30-OCT-2003.

24-APR-2002; 2002US-00132067.

24-APR-2002; 2002US-00132067.

(LALA-) LOS ALAMOS NAT LAB.  
(REGC ) UNIV CALIFORNIA.

Bradbury AM, Zeytun A, Waldo GS;

WPI; 2004-154325/15.

P-PSDB; AD136421.

PT Novel binding ligand with intrinsic fluorescence and comprising  
PT fluorescent protein having heterologous binding sites, useful for  
PT detecting target molecule.

PS Example 6; SEQ ID NO 3; 23pp; English.

CC The invention relates to binding ligands (fluorobodies) with intrinsic  
CC fluorescence, which comprises green fluorescent protein (GFP) having  
CC heterologous binding sites. The binding ligand is useful for detecting  
CC the target molecule and is efficiently detects the target molecule. The  
CC present sequence is Discosoma sp. red fluorescent protein (RED) DNA used  
CC in the exemplification of the invention.

XX SQ Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.3e-140	Length:	678
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-006-922A-12 (1-225) x AD136420 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
DB	1	ATGAGGCTTCCAAAGATGTTATCAAGGAGTTTCATGAGTTTAAAGTTCCGATGGAAGCA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACGGTCAATGGGACGAGTTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC	120
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACAATACCGTAAAGCTTAAAGTAAACCAAGGGGGACCTTGGCCATTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro	80
DB	181	TTGTCAACCAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCCCTGCCACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	241	GACTATAAAAGCTGTCTATTCTCGAAGGATTTAAATGGGAAGGTCATGAACCTTGAA	300



QY

101

AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

Db

301

GACGGTGGCGTCTGTTACTGTAAACCCAGATTCCAGTTTCAGAGATGGCTGTTTCATCTAC

360

QY

121

LyseVallysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr

140

Db

361

AAGGTCAAGTTCATTGGCGGAACTTTCCTTCGATGGACCTGTATGCAAAAGAGACA

420

QY

141

MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

Db

421

ATGGGCTGGGAAGCCAGCACTGAGCGCTTGTATCCTCGTGAATGGCGGTGTAAGAAGAGAG

480

QY

161

IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle

180

Db

481

ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTAGTTGAATTCAAAAGTATT

540

QY

181

TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp

200

Db

541

TACATGGCAAGAAGCCTGTGCAAGCTAACAGGGTACTACTAATGTTGACTCCAAACTGGAT

600

QY

201

IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLysArg

220

Db

601

ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC

660

QY

221

HisHisLeuPheLeu

225

Db

661

CACCATCTGTCTCTT

675

RESULT 15

ADM97768

ID

ADM97768

standard; DNA; 678 BP.

AC

ADM97768;

DT

01-JUL-2004

(first entry)

DE

D sp red fluorescent protein coding sequence SEQ ID NO: 21.

KW

ds; gene; enzyme; sensor cell; fluorescent protein; signal transduction detection system; promoter; targeting sequence; targeted drug.

OS

Discosoma sp.

FH

Key

Location/Qualifiers

FT

CDS

1..678

/+tag= a

FT

/product= "fluorescent protein"

XX

PN

WO2004031415-A2.

XX

PD

15-APR-2004.

XX

PF

05-SEP-2003; 2003WO-US028078.

XX

PR

05-SEP-2002; 2002US-0408297P.

XX

PA

(VERT-) VERTEX PHARM INC.

PI

Whitney MA, Zeh K, Sanders PS;

XX

DR

WPI; 2004-330208/30.

XX

DR

P-PSDB; ADM97769.

PT

Developing a sensor cell, useful in determining the activity of a target gene and in developing therapeutic drugs, comprises providing cells comprising a signal transduction detection system and introducing DNA construct into cells.

PS

Disclosure; Page 167-168; 234pp; English.

CC

The present invention relates to a method of developing a sensor cell, for determining the activity of a target gene in the cell, which

CC

comprises providing a homogeneous population of cells, where each of the cells comprises a signal transduction detection system and introducing into the population of cells an isolated DNA construct comprising a promoter operatively linked to a targeting sequence. The method is useful in developing a sensor cell for determining the activity of a target gene in the cell. The sensor cell and the methods are useful in developing new and therapeutic drugs directed to the targets. The present sequence is a coding sequence shown in the exemplification of the invention.

XX

Sequence 678 BP; 205 A; 129 C; 178 G; 166 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:

2.3e-140

Score:

1214.00

Percent Similarity:

100.00%

Best Local Similarity:

100.00%

Query Match:

100.00%

DB:

12

Length:

678

Matches:

225

Conservative:

0

Mismatches:

0

Indels:

0

Gaps:

0

US-10-006-922A-12 (1-225) x ADM97768 (1-678)

QY

1

MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly

20

Db

1

ATGAGGTCTTCCAAGAATGTTATCAAGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA

60

QY

21

ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly

40

Db

61

ACGGTCAATGGGCACGAGTTTGAATAGAAGCGAAGAGAGGGAGGCCATACGAAGGC

120

QY

41

HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle

60

Db

121

CACAAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTTGCGATATT

180

QY

61

LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro

80

Db

181

TTGTCAACAACAATTCAGTATGAAGCAAGGTATATGTCAAGCACCTCGCACATACCA

240

QY

81

AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu

100

Db

241

GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTGAA

300

QY

101

AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

Db

301

GACGGTGGCGTCTGTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTAC

360

QY

121

LyseVallysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr

140

Db

361

AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGACCTGTATGCAAAAGAGACA

420

QY

141

MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

Db

421

ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTATGGCGGTGTGAAGAAGAGAG

480

QY

161

IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle

180

Db

481

ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTAGTTGAATTCAAAAGTATT

540

QY

181

TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp

200

Db

541

TACATGGCAAGAAGCCTGTGCAAGCTAACAGGGTACTACTATGTGACTCCAAACTGGAT

600

QY

201

IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLysArg

220

Db

601

ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC

660

QY

221

HisHisLeuPheLeu

225

Db

661

CACCATCTGTCTCTT

675

RESULT 16

AAA48743

ID

AAA48743

standard; cDNA; 695 BP.

XX

CC

The present invention relates to a method of developing a sensor cell, for determining the activity of a target gene in the cell, which



AC AAA48743;  
XX  
DT 19-SEP-2000 (first entry)  
XX  
DE Humanised Discosoma sp. "red" novel fluorescent protein drFP583 cDNA.  
XX  
KM Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
XX fluorescent labeling; ss.  
XX  
OS Discosoma sp; "red".  
OS Synthetic.  
XX  
PN WO200034326-A1.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US029473.  
XX  
PR 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
PI  
XX  
XX WPI; 2000-423381/36.  
DR  
XX  
XX Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
PT  
XX  
XX Claim 6; Page 75-76; 86pp; English.  
PS  
XX  
XX The present sequence is humanised drFP583 cDNA. drFP583 is a full-length  
CC cDNA encoding a novel fluorescent protein (nFP) from Discosoma sp. "red",  
CC a non-bioluminescent species of the Class Anthozoa. The wild-type drFP583  
CC nucleotide sequence was altered to optimise the codons for expression of  
CC the fluorescent protein in mammalian cells. Fluorescent proteins can be  
CC used in fluorescent labeling, a useful tool for marking a protein, cell  
CC or organism of interest. Unlike other markers used in protein labeling,  
CC such as beta-galactosidase and luciferase, fluorescent proteins do not  
CC require an exogenous cofactor or substrate. Methods involving fluorescent  
CC proteins are also less laborious and less difficult to control than the  
CC traditional methods of fluorescent labeling, where a protein of interest  
CC is purified and then covalently conjugated to a fluorophore derivative.  
CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
CC can be used as markers for gene expression and protein localization  
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
CC They may have improved properties and better suitability for larger  
CC excitations compared to prior art fluorescent proteins such as green  
CC fluorescent protein  
CC  
SQ Sequence 695 BP; 149 A; 228 C; 209 G; 109 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.39e-140 Length: 695  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x AAA48743 (1-695)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 10 ATGCGCTCCTCCAAAGACGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 69

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 70 ACCGTGAACGCGCACGAGTTCGAGATCGAGGCGAGGGCGAGGGCCCTACAGAGGC 129

QY 41 HisAsnThrValIleuLyValThrIleuGlyGlyProLeuProPheAlaTyrAspIle 60

Db 130 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGCCCCCTGCCCTTCGCTGGACATC 189

QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80  
Db 190 CTGTCCCCCAGTTCACGATACGAGTTCAGGCTTCAAGGTGTACGTGAAGCACCCCGCATCCTCC 249

QY 81 AspTyrLyHisLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
Db 250 GACTACAAGAAGCTGTCTTCCCGGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCAG 309

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 310 GACGGCGCGTGTGACCGGTGACCCAGACTCTCCCTGCAGAGACGGCTGTCATCTAC 369

QY 121 LysValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySerThr 140  
Db 370 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAGAAGACC 429

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLyGlyGlu 160  
Db 430 ATGGCGTGGAGGCTTCACCGAGCGCTGTACCCCGGACGCGCGTGTGAAGGGCGAG 489

QY 161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
Db 490 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACTACCTGTGAGTTCAGTCCATC 549

QY 181 TyrMetAlaLySerProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200  
Db 550 TACATGCCAAGAAGCCCGTGCAGCTGCCCGCTACTACTACGTGACTCCAACTGGAC 609

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 610 ATCACCTCCCAACAAGAGACTACACCATCTGTGAAGACGAGCGCACCGAGGGCCGC 669

QY 221 HisHisLeuPheLeu 225  
Db 670 CACCACCTGTTCTTG 684

RESULT 17  
AAL47952  
ID AAL47952 standard; DNA; 859 BP.  
XX  
AC AAL47952;  
XX  
DT 26-SEP-2002 (first entry)  
XX  
DE Discosoma red fluorescent protein coding sequence.  
XX  
XX Yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;  
KM modified yeast strain; environmental pollution; gene; ds.  
XX  
XX Discosoma sp.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 54..731  
FT /\*tag= a  
FT /product= "red fluorescent protein"  
XX  
PN DE1061872-A1.  
XX  
XX 20-JUN-2002.  
XX  
XX 12-DEC-2000; 2000DE-01061872.  
PF  
XX  
PR 12-DEC-2000; 2000DE-01061872.  
XX  
XX (LICH/) LICHTENBERG-FRATZ H.  
PA  
XX  
XX Lichtenberg-Frate H;  
PI  
XX  
XX WPI; 2002-539633/58.  
DR  
DR P-PSDB; AAO18270.



XX Modified yeast strain, useful for detecting toxic compounds in  
PT environment, contains integrated cassettes responsive to genotoxic and  
PT cytotoxic compounds.  
XX  
PS Disclosure; Page 20-21; 34pp; German.

CC The present invention relates to a modified yeast strain that contains,  
CC integrated stably and functionally in its genome, a genotoxicity cassette  
CC and a cytotoxicity cassette, each comprising a promoter and reporter  
CC gene, both of which are different in the two cassettes. The modified  
CC yeast strain is used to detect environmental pollution, especially  
CC genotoxic and/or cytotoxic substances in complex environmental  
CC contaminants, especially organic compounds, but also (non-)ionizing  
CC radiation and chemical carcinogens. Particular applications are in  
CC monitoring (waste) water (e.g. as an early warning system), medical  
CC toxicology screening and for industrial process control. The present  
CC sequence is a marker gene suitable for use in the cassettes of the  
CC present invention

XX  
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.22e-140 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: / Indels: 0  
DB: Gaps: 0

US-10-006-922A-12 (1-225) x AAL47952 (1-859)

QY 1 MetArgSerSerLySBsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 54 ATGAGGTCTTCCAAGAATGTTATCAAGAGTTTCATGAGGTTAAGCTTCGCATGGAAGCA 113

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCACGAGTTGAATAGAAGCGAAGAGAGGGAGCCCATACGAAGGC 173

QY 41 HisAsnThrValLySLeuLySValThrLySGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAGCTTAAGTAACCAAGGGGGACCTTGGCCATTGGCTTGGAATATT 233

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCACAATTTCAGTATGGAAGCAAGGTAATATGTCACAGCACCTGCCGACATACCA 293

QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 294 GACTATATAAAAGCTGTCAATTTCCGGAAGATTAAATGGGAAAGGTCATGAACCTTGAA 353

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGTGGCGTCTTACTGTAACCCACGATTCAGTTTGACAGATGGCTGTTTCATCTAC 413

QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 414 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGGAACCTGTTATGCAAAAGAGACA 473

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValleuLySGlyGlu 160  
DB 474 ATGGGCTGGGAAGCCACGACGTAGCGCTTGTATCCTCGTATGGCGGTGTGAAGAGAG 533

QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
DB 534 ATTCATAAGGCTCTGAAGCTGAAAAGACGGTGTTCATTACCTAGTTGAATTCAAAAGTATT 593

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 594 TACATGGCAAGAGACCTGTGCAGCTACAGGTACTACTATGTTGACTCCAAACTGGAT 653

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 654 ATAAACAAGCCACAACGAGACTATACATCGTTGACGAGTATGAAGAACCAGAGGACGC 713  
QY 221 HisHisLeuPheLeu 225  
DB 714 CACCATCTGTCTCCTT 728

RESULT 18  
ADY51715  
ID ADY51715 standard; DNA; 859 BP.  
XX  
AC ADY51715;  
XX  
DT 05-MAY-2005 (first entry)  
XX  
DE Discosoma sp. red fluorescent protein (RFP) DNA Seq 11.  
XX  
KW fluorescence; mutagenesis; red fluorescent protein; gene; ds;  
KW protein interaction.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 54..731  
FT /\*tag= a  
FT /product= "RFP protein"  
XX  
PN WO200268605-A2.  
XX  
PD 06-SEP-2002.  
XX  
PF 26-FEB-2002; 2002WO-US006063.  
XX  
PR 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
PI Tsien RY, Baird GS, Campbell RE, Zacharias DA;  
XX  
DR WPI; 2002-713372/77.  
DR P-PSDB; ADY51716.

PT New non-oligomerizing fluorescent protein containing at least one  
PT mutation that reduces or eliminates the ability of the protein to  
PT oligomerize, useful for making better and new assays for molecular  
PT biology.

XX  
PS Disclosure; SEQ ID NO 11; 117pp; English.

XX  
CC This invention relates to a novel non-oligomerizing fluorescent protein.  
CC Specifically, it refers to the presence of at least one mutation in the  
CC fluorescent protein that reduces or eliminates the ability of the protein  
CC to oligomerize. The present invention describes fluorescent proteins and  
CC derived from naturally occurring green or red fluorescent proteins and  
CC provides a fusion protein that comprises a non-oligomerizing fluorescent  
CC protein linked to at least one protein of interest. As such, these fusion  
CC proteins can be used in methods and compositions to determine the pH of a  
CC sample, or whether the sample contains an enzyme, molecule or agent that  
CC regulates the activity of an expression control sequence. Furthermore,  
CC they may be used to identify a specific interaction of molecules, such  
CC that they are useful for improving or developing new assays in the field  
CC of molecular biology. This polynucleotide is the DNA sequence that  
CC encodes the Discosoma sp. red fluorescent protein (RFP) of the invention.

XX  
SQ Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.22e-140 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0



DB:	7	Gaps:	0
US-10-006-922A-12 (1-225) x ADY51715 (1-859)			
QY	1	MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly	20
Db	54	ATGAGGTCTTCCAGAAGATGTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATCGAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC	173
QY	41	HisAsnThrValIysLeuIysValThrIysGlyGlyProIeuProPheAlaTrpAspIle	60
Db	174	CACAAATACCGTAAAGCTTAAAGSTAAACCAAGGGGAGCTTGGCATTGCTTGGAATATT	233
QY	61	IeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	234	TTGTCAACCAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCCCTGCCGACATACCA	293
QY	81	AspTyrIlyIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAGCTGTCAATTCTCGAAGGATTTAAATGGGAAAGGTCATGAACTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerIeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTCACTAC	413
QY	121	IysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr	140
Db	414	AAGGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTTATGCAGAAAGAGACA	473
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCTCGTATGGCGTGTGAAGAGAGAG	533
QY	161	IleHisIysAlaLeuIysLeuIysAspGlyGlyHisIleTyrLeuValGluPheIysSerIle	180
Db	534	ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTTAAGTTCAAAAGTATT	593
QY	181	TyrMetAlaIysIysProValGlnIeuProGlyTyrTyrTyrValAspSerIysIeuAsp	200
Db	594	TACATGGCAAGAAGCCTGTGACGTAACAGGTACTATGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGACGC	713
QY	221	HisHisIeuPheIeu	225
Db	714	CACCATCTGTCTCTT	728
RESULT 19			
AAD53432			
ID	AAD53432 standard; DNA; 859 BP.		
XX	AAD53432;		
AC	28-MAY-2003 (first entry)		
XX	Discosoma species red fluorescent protein (RFP) encoding DNA.		
DE	Discosoma species red fluorescent protein (RFP) encoding DNA.		
XX	Phosphorylation indicator; fluorescent protein; detection; phosphatase;		
KW	kinase; red fluorescent protein; RFP; gene; ds.		
XX	Discosoma sp.		
OS	Discosoma sp.		
XX	Key		
FM	Location/Qualifiers		
FT	54..731		
PT	/*tag= a		
FT	/product= "Discosoma sp. red fluorescent protein (RFP)"		
XX			
PN	WO200295058-A2.		

XX	28-NOV-2002.	
PD	24-MAY-2002; 2002WO-US016955.	
XX	24-MAY-2001; 2001US-00865291.	
PR	(REGC ) UNIV CALIFORNIA.	
XX	Tsien RY, Ting AY, Zhang J;	
XX	WPI; 2003-148474/14.	
DR	P-PSDB; AAE34962.	
XX		
PT	Novel chimeric phosphorylation indicators, useful for detecting	
PT	kinase/phosphatase in samples, has donor molecule, phosphorylatable	
PT	domain, phosphoaminoacid binding domain, and acceptor molecule, in	
PT	operative linkage.	
XX		
PS	Disclosure; Col 64-65; 38pp; English.	
XX		
CC	The present invention relates to chimeric phosphorylation indicators	
CC	comprising a phosphorylation polypeptide and a fluorescent protein or in	
CC	operative linkage, a donor molecule, a phosphorylatable domain, a	
CC	phosphoaminoacid binding domain (PAABD) and an acceptor molecule. The	
CC	phosphorylation indicators of the invention are useful for detecting	
CC	kinases or phosphatases in a biological sample. They are also useful in	
CC	high throughput analysis e.g. for detecting a kinase inhibitor or	
CC	phosphatase inhibitor. The present sequence is Discosoma species red	
CC	fluorescent protein (RFP) encoding DNA used in the invention	
XX		
SQ	Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;	
Alignment Scores:		
Pred. No.:	3.22e-140	Length: 859
Score:	1214.00	Matches: 225
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	8	Gaps: 0
US-10-006-922A-12 (1-225) x AAD53432 (1-859)		
QY	1 MetArgSerSerIysAsnValIleIleIysGluPheMetArgPheIysValArgMetGluGly	20
Db	54 ATGAGGTCTTCCAGAAGATGTATCAAGGAGTTCATGAGGTTTAAGGTTCCGATCGAAGA	113
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	114 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGAGGGGAGGCCATACGAAGGC	173
QY	41 HisAsnThrValIysLeuIysValThrIysGlyGlyProIeuProPheAlaTrpAspIle	60
Db	174 CACATAACCGTAAAGCTTAAAGGTAAACCAAGGGGAGCTTGGCCATTGTGGGATATT	233
QY	61 IeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	234 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293
QY	81 AspTyrIlyIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	294 GACTATAAAAGCTGTCAATTCTCGAAGGATTTAAATGGGAAAGGTCATGAACTTTGAA	353
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerIeuGlnAspGlyCysPheIleTyr	120
Db	354 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGCAAGGATGGCTGTTCATCTAC	413
QY	121 IysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr	140
Db	414 AAGGTCAAGTTCATTGGCGTGAACCTTCTCCGATGACCTGTTATGCAGAAAGAGACA	473
QY	141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuIysGlyGlu	160



Db	474	ATGGGCTGGAAAGCCAGCACTAGACGCTTGTATCTCTCGTGAATGCGCTGTGAAAGAGAG	533
Qy	161	ILEHISLYSALALEULYSLEULYSAAPGIGLYHISLYRLEUVALGLUPHELYSERILE	180
Db	534	ATTGATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT	593
Qy	181	TYRMETALALYSLYSPROVALGINLEUPROGLYTYRTRYRTRYRVALASPSELYSLEUASP	200
Db	594	TACATGGCAAGAACCTGTGACAGCTACCAAGGGTACTACTATGTGACTCCAAACTGGAT	653
Qy	201	ILETHRSEHISASNGLUASPTYRTHRILEVALGLUGINTYRGLUARGTHRGLUGLYARG	220
Db	654	ATAACAAGCCACAAAGCAAGACTATTAACAATCGTTGAGCAGTATGAAAGAACCGAGGACGC	713
Qy	221	HISHSLEUPHELEU	225
Db	714	CACCATCTGTCTCTT	728
RESULT 20			
ID	AAD61969	standard; cDNA; 859 BP.	
AC	AAD61969;		
XX	15-JAN-2004	(first entry)	
XX			
DE	Discosoma sp. red fluorescent protein (RFP) cDNA.		
XX			
KW	Fluorescent protein; resonance energy transfer; pH; detection;		
KW	red fluorescent protein; RFP; gene; ss.		
XX			
OS	Discosoma sp.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	54..731	
FT		/*tag= a	
FT		/product= "Discosoma sp. red fluorescent protein"	
XX			
PN	US2003170911-A1.		
XX			
PD	11-SEP-2003.		
XX			
PF	26-FEB-2001; 2001US-00794308.		
XX			
PR	26-FEB-2001; 2001US-00794308.		
XX			
PA	(TSIE/) TSIEH R Y.		
PA	(ZACH/) ZACHARIAS D A.		
PA	(BAIR/) BAIRD G S.		
XX			
PI	Tsien RY, Zacharias DA, Baird GS;		
XX			
DR	WPI; 2003-802418/75.		
DR	P-PSDB; ABW00918.		
XX			
PT	Fluorescent proteins containing a mutation that reduces or eliminates its		
PT	ability to oligomerize which gives more reliable fluorescence resonance		
PT	energy transfer results and are useful to detect molecule interaction,		
PT	enzymes, or sample pH.		
XX			
PS	Disclosure; Page 29-30; Opp; English.		
XX			
CC	The invention relates to a non-oligomerizing fluorescent protein		
CC	containing a mutation that reduces or eliminates its ability to		
CC	oligomerize. The fluorescent protein gives more reliable fluorescence		
CC	resonance energy transfer results and are useful to detect molecule		
CC	interaction, enzymes, or sample pH. These are also used to identify		
CC	agents or conditions that regulate expression of control sequences. The		
CC	present sequence is Discosoma sp. red fluorescent protein (RFP) cDNA		
XX			
SO	Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;		

Pred. No.:	3.22e-140	Length:	859
Score:	1214.00	Matches:	225
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922A-12 (1-225) x AAD61969 (1-859)

QY		1 MetArgSerSerIyAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db		54 ATGAGGTCTTCCAAGAATGTTATCAAGAGATTTCATGAGGTTTAAGTTCCGATGAAGGA	113
QY		21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugLy	40
Db		114 ACGGTCAATGGGCACGAGTTTGAAATAGAAGGCCAAGAGAGAGGGAGGCCATACGAAGGC	173
QY		41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db		174 CACAATACCCTAAAGCTTAAGGTAAACCAAGGGGGGACCCTTGGCATTGCTTGGGATATT	233
QY		61 LeuSerProGlnPheGlnTyrgLysSerLysValTyrrValLysHisProAlaAspIlePro	80
Db		234 TTGTCAACCACAATTTCAGTAGAGAACCAAGGTATATGTCAAGCACCTGCCGACATACCA	293
QY		81 AspTyrrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db		294 GACTATAAAAAGCTGTCATTTCCGAAGGATTTAAATGGAAAGGTCATGAACCTTGAA	353
QY		101 AspGlyGlyValValThrValThrcGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db		354 GACGGTGGCGTCGTACTGTAAACCAAGGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC	413
QY		121 LysValLysPheIleGlyValAsnDheProSerAspGlyProValMetGlnLysLysThr	140
Db		414 AAGGTCAAAGTTCATTTGGCGTGAACCTTCTCCGATGGAACCTGTTATGC AAAAGAGACA	473
QY		141 MecGlyTrpGluAlaSerThrcGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db		474 ATGGGCTGGGAAGCCAGCACGTAGCGCTTGATCCTCGTAGTGCGGTGTAAGAGAGAG	533
QY		161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db		534 ATTTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCATTACCTAGTTGAATCAAAAGTATT	593
QY		181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db		594 TACATGGCAAGAACCTGTGCAGCTAACGAGGTACTACTATGTGTGACTCCAACTGGAT	653
QY		201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrcGluGlyArg	220
Db		654 ATAACAGACCAACAAGACTATACATCGTTGAGCAGTATGAAGAAGACGAGGAGCGC	713
QY		221 HisHisLeuPheLeu 225	
Db		714 CACCATCTGTCTCTT 728	
RESULT 21			
ID	ADX26533	standard; DNA; 859 BP.	
XX	ADX26533;		
XX	AC		
DT	21-APR-2005	(first entry)	
DE	Discosoma DsRed RFP protein coding sequence, seq id 11.		
KW	Phosphorylation; detection; red fluorescent protein; RFP; gene; ds.		
OS	Discosoma sp.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	54..731	



FT /tag= a  
PT /product= "DfRed red fluorescent protein"  
XX  
XX US2005026234-A1.  
XX PD 03-FEB-2005.  
XX  
XX 28-MAY-2004; 2004US-00857622.  
XX  
XX 31-JAN-1996; 96US-00594575.  
XX PR 31-JAN-1997; 97US-00792553.  
XX PR 13-SEP-1999; 99US-00396003.  
XX PR 24-MAY-2001; 2001US-00865291.  
XX  
XX (VIOL/) VIOLIN J D.  
XX PA (NEWT/) NEWTON A C.  
XX PA (TSIE/) TSIEEN R Y.  
XX PA (ZHAN/) ZHANG J.  
XX  
XX Violin JD, Newton AC, Tsien RY, Zhang J;  
XX WPI; 2005-141377/15.  
XX DR P-PSDB; ADX26534.  
XX  
XX New chimeric phosphorylation indicator comprises a first fluorescent  
PT protein, phosphoaminoacid binding domain with FHA2 sequence and protein  
PT kinase C-phosphorylatable domain, useful for detecting protein kinase C  
PT or phosphates.  
XX  
XX Disclosure; SEQ ID NO 11; 68bp; English.  
XX  
XX The invention relates to a chimeric phosphorylation indicator (CPI) (I)  
CC comprising in operative linkage, a first fluorescent protein, a  
CC phosphoaminoacid binding domain with an FHA2 sequence defined in the  
CC specification, a protein kinase C (PKC)-phosphorylatable domain, and a  
CC second fluorescent protein. The first and second fluorescent proteins are  
CC chosen from green fluorescent proteins (GFPs) and red fluorescent  
CC proteins (RFPs). The first and second fluorescent proteins exhibit a  
CC detectable resonance energy transfer, when the first fluorescent protein  
CC is excited. The PKC-phosphorylatable domain and phosphoaminoacid binding  
CC domain do not substantially emit light to excite the second fluorescent  
CC protein. (I) is useful for detecting a protein kinase C (PKC) or  
CC phosphates in a sample (such as biological sample e.g. cell, tissue  
CC sample, or their extracts). (I) is useful for detecting a kinase  
CC inhibitor or phosphatase inhibitor. The method of the invention is  
CC adapted to high throughput analysis. The current sequence represents the  
CC coding sequence of Discosoma DfRed RFP.  
XX  
XX Sequence 859 BP; 263 A; 163 C; 228 G; 205 T; 0 U; 0 Other;  
SQ  
  
Alignment Scores:  
Pred. No.: 3.22e-140 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
  
US-10-006-922A-12 (1-225) x ADX26533 (1-859)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 54 ATGAGGTCTTCCAGAATGTTATCAAGAGTTCATGAGGTTTAAAGTTGCGATGGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGAAGGAGAGGGGACCATACGAAGGC 173  
QY 41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAAATACCGTAAAGCTTAAGGTAAACCAAGGGGAGACCTTTGCCATTGCTTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80

Db 234 TTGTCACCACCAATTTCAGTATGGAAGCAAGGTATATGTCAAAGCACCCTGCCGACATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAGCTGTCATTCTTCTGAAAGATTAAATGGGAAAAGGTCATGAACTTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCTGTAAGTGAACCAAGATTCCAGTTTGCAAGATGGCTGTTCACTTAC 413  
QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 414 AAGGTCAAGTTCATGGCGTGAACCTTCCATGATGGAAGCTTATGCAAAAGAGACA 473  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGATGGCGTGTGAAGAGAG 533  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 594 TACATGGCAAGAAGCCTGTGACAGCTACAGGCTACTATATGTTGACTCCAAACTGGAT 653  
QY 201 IleThrSerHisAsnGlyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 654 ATAACAAGCCACAACAGAGACTATACAATCGTTGAGCAGTATGAAGAAGACGAGGAGCGC 713  
QY 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCCTT 728  
  
RESULT 22  
ACA62995  
ID ACA62995 standard; DNA; 3311 BP.  
XX  
XX ACA62995;  
AC  
XX  
DT 23-SEP-2003 (first entry)  
XX  
XX Plasmid DNA containing coding sequence for RFP.  
DE  
XX Fluorescently-tagged enzyme; substrate; cell population; GFP;  
KW quantification of enzymatic activity; green fluorescent protein;  
KW red fluorescent protein; RFP; enhanced green fluorescent protein; EGFP;  
KW enzymatic process; cellular enzyme; chemotherapeutic drug;  
KW multidrug resistance; MDR; ds.  
XX  
XX Aequorea victoria.  
OS  
OS Synthetic.  
XX  
XX US2003049597-A1.  
PN  
XX  
XX 13-MAR-2003.  
XX  
XX PD  
XX  
XX 01-MAR-2001; 2001US-00797496.  
PF  
XX  
XX  
XX 01-MAR-2001; 2001US-00797496.  
PR  
XX  
XX (SIMO/) SIMON S M.  
PA (CHEN/) CHEN Y.  
XX  
XX Simon SM, Chen Y;  
PI  
XX  
XX WPI; 2003-555145/52.  
DR  
XX  
XX Simultaneously quantifying in situ the relationship between an enzyme and  
PT its substrate for study of enzymatic processes at a cellular level;  
PT comprises optical measurements on cells expressing the fluorescently  
PT tagged enzyme.



XX Disclosure; Page 8-9; 21pp; English.  
XX  
CC The present invention relates to a method for simultaneously quantifying  
CC its substrate in a population of cells. A population of cells expressing  
CC (E) is created in which different cells contain different amounts of (E).  
CC The cells are incubated with a substrate, and the concentration and  
CC enzymatic activity of (E) in each cell are simultaneously quantified by  
CC optical means. The fluorescently-tagged enzyme is produced by linking the  
CC enzyme with Aequorea victoria green fluorescent protein (GFP), or its  
CC variants or derivatives such as red fluorescent protein (RFP) and  
CC enhanced green fluorescent protein (EGFP). The method of the invention is  
CC useful for simultaneously quantifying in situ the relationship between an  
CC enzyme and its substrate in a cell. The method is useful for the study of  
CC enzymatic processes at the cellular level, and especially for examining  
CC the interactions of cellular enzymes with chemotherapeutic drugs, e.g.  
CC for elucidating mechanisms of multidrug resistance (MDR). The present  
CC sequence represents a plasmid containing the coding sequence for RFP  
XX  
SQ Sequence 3311 BP; 858 A; 778 C; 860 G; 815 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.15e-139 Length: 3311  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x ACA62995 (1-3311)

QY 1 MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGly 20  
DB 289 ATGAGGCTTCCAAAGATGTATCAAGAGCTTCATGAGGTTAAGGTTCCGATGGAAGGA 348  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 349 ACGGTCAATGGGACAGATTGAAATAGAGCGAAGAGAGGGGAGCCATACGAAGGC 408  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 409 CACAATACCGTAAGCTTAAGTAACCAAGGGGGACCTTGCCATTGCTGGGATATT 468  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 469 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCGCGACATACCA 528  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 529 GACTATAAAGGCTGTCTATTCTTGAAGGATTAAATGGGAAAGGCTCATGAACCTTGAA 588  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 589 GACGGTGGCGTCTTAAGTGAACCAAGATTCAGTTGCAAGATGGCTGTTTCATCTAC 648  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 649 AAGGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGAGACCTGTATGCAAAAGAGACA 708  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 709 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCGTGATGCGCGTGTGAAGAAGAGAG 768  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 769 ATTCAATAGGCTCTGAAGCTGAAGAAGCGGTGTCTATTACCTAGTTGAATCAAAAGTATT 828  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 829 TACATGGCAAGAGCCTGTGTGACGCTACCAAGGTAAGTATGTTGACTCCAAACTGGAT 888  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 889 ATAACAAGCCACAACGAAGACTATACATCTGTGAGCAGTATGAAAGAACCGAGGACGC 948  
QY 221 HisHisLeuPheLeu 225  
DB 949 CACCACTCTGTTCCTT 963

RESULT 23  
ID ADC24134  
ID ADC24134 standard; DNA; 681 BP.

AC ADC24134;  
DT 18-DEC-2003 (first entry)

DE Discosoma wild-type red fluorescent protein DNA #2.  
XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
KW protein phosphatase; ion indicator; ds.

XX Discosoma.

XX US2003059835-A1.  
XX 27-MAR-2003.

PF 10-APR-2002; 2002US-00121258.

PR 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.

XX (TSIE/) TSIE N Y.  
PA (CAMP/) CAMPBELL R E.

XX Tsien RY, Campbell RE;

DR WPI; 2003-743764/70.  
DR P-PSDB; ADC24126.

XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.

XX Disclosure; SEQ ID NO 9; 67pp; English.

XX The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca<sup>2+</sup>, Zn<sup>2+</sup>, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is



CC useful for identifying a region or condition that regulates the activity  
 CC of an expression control sequence. This sequence represents a modified  
 CC Discosoma wild-type red fluorescent protein coding region with humanised  
 CC codon usage.

XX  
 SQ Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	7.26e-140	Length:	681
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922a-12 (1-225) x ADC24134 (1-681)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
Db	4	GTGCGCTCCTCCAAAGACGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC	63
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	64	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGC	123
QY	41	HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
Db	124	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTGGGACATC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
Db	184	CTGTCCCCCAAGTTCACAGTACGGCTCCAAAGGTGATCGTGAAGCACCCCGCGACATCCC	243
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu	100
Db	244	GACTCAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG	303
QY	101	AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	304	GACGGCGCGGTGTGACCGTGACCAAGACTCCTCCCTGCGAGGACGGGCTTCATTCAC	363
QY	121	LySValLySLeuIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140
Db	364	AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCGTAATGCAAGAAAGACC	423
QY	141	MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160
Db	424	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCCGCGACGGCTGTGAAGGGCGAG	483
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180
Db	484	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGCTGAGTTCAAGTTCATC	543
QY	181	TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
Db	544	TACATGGCCAAAGAGCCCGTGACGTGCCCGGCTACTACTAGTGAAGTGAAGTGGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	604	ATCACTCCCAACAAGAGACTACATCGTGAGCAGTACGAGCGCACCGAGGGCGGC	663
QY	221	HisHisLeuPheLeu	225
Db	664	CACCACTGTTCCTG	678
RESULT 24			
ADL46205			
ID	ADL46205 standard; DNA; 681 BP.		
XX			
AC	ADL46205;		
XX			
DT	20-MAY-2004 (first entry)		
.XX			

DE Human codon optimised Discosoma red fluorescent protein (DsRed) DNA.  
 XX  
 XX ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
 KW oligomerization; tetramerization; immunoassay; hybridization assay.  
 XX  
 XX  
 OS Discosoma sp.

XX  
 XX WO2003086446-A1.  
 PN  
 XX 23-OCT-2003.

XX  
 PD 09-APR-2003; 2003WO-US010879.

XX  
 PF 10-APR-2002; 2002US-00121258.  
 PR 29-JUL-2002; 2002US-00209208.

XX  
 PA (REGC ) UNIV CALIFORNIA.

XX  
 PI Tsien RY, Campbell RE, Baird GS;

XX  
 DR WPI; 2003-845265/78.

XX  
 PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
 PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
 PT in molecular biology, e.g. in immunoassays or in tracking protein  
 PT movement in cells.

XX  
 PS Disclosure; SEQ ID NO 3; 166pp; English.

XX  
 CC The invention relates to a polynucleotide sequence encoding a Discosoma  
 CC red fluorescent protein (DsRed) variant having a reduced propensity to  
 CC oligomerize. The protein variant comprises one or more amino acid  
 CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
 CC sequence, where the substitutions result in reduced propensity of the  
 CC DsRed variant to form tetramers and where the variant displays detectable  
 CC fluorescence of at least one red wavelength. The composition and methods  
 CC are useful in producing red fluorescent proteins having reduced  
 CC propensity for oligomerization, especially tetramerization. The protein  
 CC may be used in molecular biology and in other scientific applications,  
 CC such as in immunoassays or hybridization assays, or in tracking the  
 CC movement of proteins in cells. This sequence corresponds to the DsRed  
 CC coding sequence with codons optimised for human expression.

SQ Sequence 681 BP; 146 A; 223 C; 205 G; 107 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	7.26e-140	Length:	681
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	11	Gaps:	0

US-10-006-922a-12 (1-225) x ADL46205 (1-681)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
Db	4	GTGCGCTCCTCCAAAGACGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC	63
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	64	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC	123
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
Db	124	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTGGGACATC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
Db	184	CTGTCCCCCAAGTTCACAGTACGGCTCCAAAGGTGATCGTGAAGCACCCCGCGACATCCC	243
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu	100







XX 25-MAR-2003 (first entry)  
DT Mammalian codon optimised Discosoma red fluorescent protein DNA.  
XX  
DE Mammalian codon optimised Discosoma red fluorescent protein; Discosoma;  
XX Mammalian codon optimised Discosoma red fluorescent protein; gene; ds.  
KW red fluorescent protein; directed evolution; gene; ds.  
XX  
OS Discosoma sp.  
OS Mammalia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..723  
FT /tag= a  
FT /product= "Mammalian codon optimised Discosoma red  
FT fluorescent protein"  
FT /transl\_except= {pos:616..618,aa:His)  
XX  
XX WO200294992-A2.  
XX  
XX 28-NOV-2002.  
PD  
XX 20-MAY-2002; 2002WO-US015968.  
PF  
XX 18-MAY-2001; 2001US-0291871P.  
PR  
XX (RIGE-) RIGEL PHARM INC.  
PA  
XX  
XX Peel1e B;  
PI  
XX  
DR WPI; 2003-120798/11.  
DR P-PSDB; ABP56678.  
XX  
XX New Discosoma red fluorescent protein, useful for functional screens as a  
PT reporter for gene transcription, for target characterization and  
PT localization of fusion proteins, or for scaffolds for protein and peptide  
PT libraries.  
XX  
XX Example 1; Fig 1; 22pp; English.  
PS  
XX  
XX The present invention describes an isolated Discosoma red fluorescent  
CC protein (I) comprising a 241 residue amino acid sequence (see ABP56678,  
CC S1), with one or more point mutations at amino acid position N24, F125,  
CC K164, or M183. Also described: (1) a fusion protein comprising (I); (2)  
CC an isolated nucleic acid encoding (I); (3) a vector comprising the  
CC nucleic acid of (2); (4) a host cell comprising the vector; (5) a  
CC retroviral cDNA expression library comprising the nucleic acid of (2), or  
CC encoding (I); and (6) methods of making a fluorescent variant. The  
CC Discosoma red fluorescent proteins are useful for functional screens as a  
CC reporter for gene transcription (e.g. as a fusion protein), for target  
CC characterisation and localisation of fusion proteins, or for scaffolds  
CC for protein and peptide libraries. The fluorescent proteins can also be  
CC used as selectable markers or reporter molecules for a variety of  
CC bioassays, including methods that use fluorescence activated cell sorting  
CC (FACS) as a selection mechanism. The method of directed protein evolution  
CC is useful for obtaining improved variants of red fluorescent protein. The  
CC variants of Discosoma red fluorescent protein have greatly improved  
CC brightness, expression, and/or folding kinetics as compared to wild type  
CC or a codon optimised variant. The present sequence encodes a mammalian  
CC codon optimised Discosoma red fluorescent protein, which is used in an  
CC example from the present invention  
XX  
XX Sequence 723 BP; 161 A; 233 C; 221 G; 108 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	7.9e-140
Score:	1210.00
Percent Similarity:	100.00%
Best Local Similarity:	99.56%
Query Match:	8
DB:	8
	Gaps:
	Mismatches:
	Conservative:
	Matches:
	Length:
	723
	224

US-10-006-922A-12 (1-225) x ABZ22476 (1-723)

QY	1	MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGlu	20
Db	4	GTGGCTCTCCCAAGAACGTATCAAGAGTTTATGCGCTTCAAGGTGCGCATGGAGGGC	63
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	64	ACCGTGAACGGCCACGAGTTTCAGATGACGGCGAGGGCGAGGGCCGCCCTACGAGGGC	123
QY	41	HisAsnThrValIysLeuLysValThrIysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	124	CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCTGGGACATC	183
QY	61	LeuSerProGlnPheGlnTyrGlySerIysValTyrValLysHisProAlaAspIlePro	80
Db	184	CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGTACGTGAAGCACCCCGCAGATCCCC	243
QY	81	AspTyrIleLysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	244	GACTACAAGAACTGTCTCTCCCGAGGCTTCAAGTGGAGCGCGTGTAACTTCGAG	303
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	304	GACGCGCGCGTGGTGACCGTGACCCAGAACTCTCCCTGCAGACGGCTGTTCATCTAC	363
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	364	AAGGTGAAGTTCATCGGCGTGAACTTCCCTCCGACGGCCCGTAATGCAGAAAGAAC	423
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	424	ATGGGCTGGGAGGCCCTCCACCAGCGCCTGTACCCCCCGACGGCGGTGTGAAGGGCGAG	483
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	484	ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCCACTACCTGTGTGAGTTCAAGAGTATC	543
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	544	TACATGGCCCAAGAACCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAAGCTGGAC	603
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	604	ATCACCTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC	663
QY	221	HisHisLeuPheLeu	225
Db	664	CACCACCTGTTCCTG	678

	RESULT 27
ID	ADL18131
ID	ADL18131 standard; DNA; 1638 BP.
XX	
AC	ADL18131;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	RFP:PS(NIa protease):AtOEP7:GFP fusion protein DNA SEQ ID NO:51.
XX	
KW	chimeric protein; signal protein; trafficking signal targeting;
KX	proteolytic cleavage site; protease; protease inhibitor; gene; ds
XX	
OS	Arabidopsis thaliana.
OS	Tobacco vein mottling virus.
OS	Synthetic.
XX	
PN	WO2003014381-A1.
XX	
PD	20-FEB-2003.
XX	
PF	08-AUG-2002; 2002WO-KR001515.
XX	



PR 10-AUG-2001, 2001KR-00048123.  
XX  
PA (AHRA-) AHRAM BIOSYSTEMS INC.  
XX  
PI Hwang I, Kim DH, Lee YJ;  
XX  
DR WPI, 2003-256596/25.  
DR P-PSDB; ADL18132.  
XX  
PT New chimeric protein, useful for detecting protease inhibitors inside the  
PT cell or tissue.  
XX  
PS Example 2; SEQ ID NO 51; 214pp; English.  
XX  
CC The present invention describes a chimeric protein comprising at least  
CC one signal protein that has a trafficking signal targeting to a  
CC subcellular organelle and at least one proteolytic cleavage site for a  
CC protease. The chimeric protein is constructed, so that: (a) the  
CC trafficking signals of all the signal proteins are inactivated by linking  
CC the proteolytic site or a signal masking protein through the proteolytic  
CC site to the N-or C- terminus of the signal proteins, and so the chimeric  
CC protein is present in cytosol; (b) the trafficking signal of at least one  
CC signal protein is activated when the proteolytic cleavage site is cleaved  
CC by the protease, and as a result at least one fragment protein that  
CC includes the activated signal protein is a transported to a subcellular  
CC organelle; and (c) the chimeric protein is labelled with at least one  
CC fluorescent protein and the position and intensity distribution of the  
CC fluorescent label signal in the cell is altered depending on the cleavage  
CC by the protease. Also described: (1) a recombinant gene comprising a  
CC nucleic acid sequence encoding the chimeric protein which is constructed  
CC to express the chimeric protein in a cell; (2) a cell transformed with  
CC the recombinant gene or vector; (3) analysing the activity of a protease  
CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
CC detecting a protease inside a cell; (6) a nucleic acid comprising the  
CC sequence encoding the chimeric protein for detecting protease activity in  
CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
CC a protease inside a cell comprising the chimeric protein or the vector;  
CC (9) detecting a protease inside a cell or tissue; and (10) detecting a  
CC protease inhibitor in vivo. The chimeric protein is useful for detecting  
CC protease inhibitors inside the cell or tissue. The present sequence  
CC encodes a fusion protein, which is used in the exemplification of the  
CC present invention.  
XX  
XX  
SQ Sequence 1638 BP; 445 A; 450 C; 419 G; 324 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.5e-139 Length: 1638  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADL18131 (1-1638)

QY 1 MetArgSerSerLybAsnValIleLysGluPheMetArgPheLybValArgMetGluGly 20  
DB :::  
4 GTGCGCTCCTCCAGAGACGTATCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCAGAGTTGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAGGTGAAGTGAACCAAGGGCGGCCCCCTTGCCCTGCGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAAGTTCAGTACGGCTCCAAGGTGTACGTGAAGACACCCCGCGACATCCCC 243  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100

DB 244 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 304 GACGGCGGGCGTGTGACCGGTGACCCAGACTCTCCCTCGAAGACGGCTGTTCACTAC 363  
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLeuThr 140  
DB 364 AAGGTGAAGTTCATCGCGGTGAATTCCTCCCGACGGCCCCGTAATGAGAAGAAGACC 423  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160  
DB 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCCCGACGGCGTGCTGAAGGCGAG 483  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGACGGCGCCACTACTACTACTGAGACTCCAAGCTGAC 543  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
DB 544 TACATGCCCCAAGAAGCCCGTGCAGCTGCCGCTACTACTACTGAGACTCCAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 604 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAATGACGACGACGAGGCGCGC 663  
QY 221 HisHisLeuPheLeu 225  
DB 664 CACCACTGTCTCTG 678

RESULT 28

ADL18155 standard; DNA; 1647 BP.

ADL18155;

06-MAY-2004 (first entry)

RFP:PS(HIV-1 protease):AtOEP7:GFP fusion protein DNA SEQ ID NO:75.

chimeric protein; signal protein; trafficking signal targeting;

proteolytic cleavage site; protease; protease inhibitor; gene; ds.

Arabidopsis thaliana.

Human immunodeficiency virus 1.

Synthetic.

WO2003014381-A1.

20-FEB-2003.

08-AUG-2002; 2002WO-KR001515.

10-AUG-2001; 2001KR-00048123.

(AHRA-) AHRAM BIOSYSTEMS INC.

Hwang I, Kim DH, Lee YJ;

WPI; 2003-256596/25.

P-PSDB; ADL18156.

New chimeric protein, useful for detecting protease inhibitors inside the

cell or tissue.

Example 3; SEQ ID NO 75; 214pp; English.

The present invention describes a chimeric protein comprising at least

one signal protein that has a trafficking signal targeting to a

subcellular organelle and at least one proteolytic cleavage site for a

protease. The chimeric protein is constructed, so that: (a) the

trafficking signals of all the signal proteins are inactivated by linking

the proteolytic site or a signal masking protein through the proteolytic



CC site to the N-or C- terminus of the signal proteins, and so the chimeric  
CC protein is present in cytosol; (b) the trafficking signal of at least one  
CC signal protein is activated when the proteolytic cleavage site is cleaved  
CC by the protease, and as a result at least one fragment protein that  
CC includes the activated signal protein is a transported to a subcellular  
CC organelle; and (c) the chimeric protein is labelled with at least one  
CC fluorescent protein and the position and intensity distribution of the  
CC fluorescent label signal in the cell is altered depending on the cleavage  
CC by the protease. Also described: (1) a recombinant gene comprising a  
CC nucleic acid sequence encoding the chimeric protein which is constructed  
CC to express the chimeric protein in a cell; (2) a cell transformed with  
CC the recombinant gene or vector; (3) analysing the activity of a protease  
CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for  
CC detecting a protease inside a cell; (6) a nucleic acid comprising the  
CC sequence encoding the chimeric protein for detecting protease activity in  
CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting  
CC a protease inside a cell comprising the chimeric protein or the vector;  
CC (9) detecting a protease inside a cell or tissue; and (10) detecting a  
CC protease inhibitor in vivo. The chimeric protein is useful for detecting  
CC protease inhibitors inside the cell or tissue. The present sequence  
CC encodes a fusion protein, which is used in the exemplification of the  
CC present invention.

XX SQ Sequence 1647 BP; 450 A; 448 C; 421 G; 328 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.52e-139 Length: 1647  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADL18155 (1-1647)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db :::  
4 GTGGCGTCTCTCAAGAACGTCATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACGGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLySLeuLySerValThrLySGlyGlyProLeuProPheAlaTyrPAspIle 60  
Db 124 CACAACACCGTGAAAGTGAGCAAGGGCGGCCCTTGGCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySerLyValTyrValLyHisProAlaAspIlePro 80  
Db 184 CTGTCCCCCAGTTCAGTACGGGCTCCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 243  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
Db 244 GACTACAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 304 GACGGCGGCGTGGTGACCGTGACCAAGACTCTCCCTGCAGAGACGGCTGCTTCATCTAC 363  
QY 121 LySValLySPheIleGlyValAsnProSerAspGlyProValMetGlnLySLeuThr 140  
Db 364 AAGGTGAAGTTCATCGCGTGAACTCCCTCCGACGGCCCGTAATGCAGAAAGAAC 423  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 424 ATGGGTGGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGGTGAGAGGGCGAG 483  
QY 161 IleHisLySAlaLeuLySLeuLyAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCCACTACCTGTGAGTTCAAGTCCATC 543  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 544 TACATGCCAAGAGCCCGTGAGCTGCCCGCTACTACTACGTGACTCCAAGCTGAC 603  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 604 ATCACCTCCACAACGAGGAGTACACCATCGTGGAGCAGTACGAGCGCACCGAGGCCGC 663  
QY 221 HisHisLeuPheLeu 225  
Db 664 CACCACCTGTCTCTG 678  
RESULT 29  
ADS75466  
ID ADS75466 standard; DNA; 2721 BP.  
XX  
AC ADS75466;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Fibrohexamerin promoter, exon 1 and intron 1 synthetic sequence SEQ ID 1.  
XX  
KW Spidroin; fibroin; silkworm; Ds-Red; Fibrohexamerin; gene; promoter; ds.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT 1..1452  
FT promoter  
FT /tag= a  
FT /label= Fibrohexamerin promoter  
FT misc\_feature 1150..2026  
FT /tag= b  
FT /note= "This region is specifically claimed in Claim 2"  
FT CAAT\_signal 1362..1365  
FT /tag= c  
FT misc\_signal 1379..1390  
FT /tag= d  
FT /note= "Binding site for transcription factors of the  
FT TATA\_signal 1420..1423  
FT /tag= e  
FT exon 1452..1525  
FT /tag= f  
FT /number= 1  
FT /note= "Fibrohexamerin exon 1, encodes the signal  
FT intron 1526..2026  
FT /tag= g  
FT /number= 1  
FT CDS /note= "Fibrohexamerin intron 1"  
FT 2041..2721  
FT /tag= h  
FT /product= "Reporter protein Ds-Red"  
XX  
PN FR2852325-A1.  
PN  
PD 17-SEP-2004.  
XX  
XX 13-MAR-2003; 2003FR-00003137.  
PF  
XX 13-MAR-2003; 2003FR-00003137.  
PR 13-MAR-2003; 2003FR-00003137.  
XX  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
PA (UYLY-) UNIV LYON 1 BERNARD CLAUDE.  
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
PI Chavancy G, Couble P, Durand B, Grenier A, Horard B, Julien E;  
PI Mauchamp B, Nony P, Prudhomme JC, Royer C;  
XX  
DR WPI; 2004-670997/66.  
XX  
XX New nucleic acid that directs expression of protein in silk glands of  
PT silkworm, useful for preparation of pharmaceutical proteins and for  
PT modifying textile properties of silk.  
XX



PS Claim 1; SEQ ID NO 1; 50pp; French.

XX The present invention relates to a novel nucleic acid (I) that directs  
CC expression of a protein of interest (II) specifically in the cells of the  
CC posterior silk-producing glands (A) of Bombyx mori. (I) comprises, in the  
CC 5' to 3' direction: a regulatory region (RR) comprising signals for  
CC regulating expression of a polynucleotide of interest (III) in the  
CC specified cells and, under control of RR, a region that encodes a  
CC modified signal peptide (msp). (I) is at least 90% identical with the  
CC 1150-2026 nucleotide (nt) region of ADS75466, in which the codon  
CC containing nt 1486-1488 encodes Ala, Ile or Leu. (I) is optionally  
CC modified by insertion of 1-4 copies of the sequence ADS75467 (nucleotides  
CC 1379-1390 of ADS75466), which is the binding site for transcription  
CC factors of the SGF1 forkhead type, functioning as activator of  
CC transcription and regulator of tissue-specific expression. (III) encodes  
CC a polypeptide, specifically spidroin (of spiders) or fibroin (from  
CC organisms of the genus Galleria). (I), also related expression cassettes  
CC and vectors, are used to prepare transgenic Bombyx that secrete (II)  
CC incorporated into silk threads, e.g. where (II) is of biomedical interest  
CC (e.g. hormones, antigens, enzymes, growth factors and receptors) or where  
CC the threads are intended for use as textiles and (II) modifies resistance  
CC to ductility or elasticity.

SQ Sequence 2721 BP; 748 A; 631 C; 606 G; 736 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.11e-139 Length: 2721  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x ADS75466 (1-2721)

OY 1 MetArgSerSerIyAsnValIleIyGluPheMetArgPheIyValArgMetGluIy 20  
DB 2044 GTGCGCTCTCTCCAAAGACGTATCATAGGAGTTTCATGCGCTCAAGTGGCATGAGGGC 2103  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluIyArgProTyrGluGly 40  
DB 2104 ACCGTGAACGGCCAGAGTTCAGATCGAGGGCGAGGGCCGCCCTTACGAGGGC 2163  
OY 41 HisAsnThrValIySleuIyValThrIySgIyGlyProIeuProPheAlaTyrAspIle 60  
DB 2164 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGGCTGGACATC 2223  
OY 61 IeuSerProGlnPheGlnTyrGlySerIyValTyrValIySHsProAlaAspIlePro 80  
DB 2224 CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGACGTGAAGCACCCCGCGACATCCCC 2283  
OY 81 AspTyrIySleuSerPheProGluGlyPheIySTrpGluArgValMetAsnPheGlu 100  
DB 2284 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCGAG 2343  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerIeuGlnAspGlyCySPheIleTyr 120  
DB 2344 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGCAAGACGGCTCTTCACTAC 2403  
OY 121 IySValIySPheIleGlyValAsnPheProSerAspGlyProValMetGlnIySlyThr 140  
DB 2404 AAGGTGAAGTTCATCGGCGTGAATTCTCCCTCCGAGGGCCCGTAAATGCAAGAAAGACC 2463  
OY 141 MetGlyTrpGluAlaSerThrGluArgIeuTyrProArgAspGlyValIeuIySgIyGlu 160  
DB 2464 ATGGGCTGGAGGCTTCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGGCGAG 2523  
OY 161 IleHisIySAlaIeuIySleuIySAspGlyGlyHisTyrIeuValGluPheIySerIle 180  
DB 2524 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCACCTACTGTGAGTTCAAGTCCATC 2583  
OY 181 TyrMetAlaIySleuProValGlnIeuProGlyTyrTyrTyrValAspSerIySleuAsp 200

DB 2584 TACATGGCCAAGAGCCCGTGACGTGCCCGCTACTACTAGTGGACTCCAAGCTGGAC 2643  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 2644 ATCACTTCCACACAGAGACTACACCATCTGTGAGCAGTACGAGCGACCGAGGGCCGC 2703  
OY 221 HisHisIeuPheIeu 225  
DB 2704 CACCACCTGTCTCTG 2718  
RESULT 30  
ADS75468  
ID ADS75468 standard; DNA; 2772 BP.  
XX  
AC ADS75468;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Fibrohexamerin promoter, exon 1 and intron 1 synthetic sequence SEQ ID 3.  
XX Spidroin; fibroin; silkworm; Ds-Red; Fibrohexamerin; gene; promoter; ds.  
XX  
OS Synthetic.  
XX  
FH Key location/Qualifiers  
FT promoter 1..1503  
FT /\*tag= a  
FT /label= Fibrohexamerin promoter  
FT CAAAT\_signal 1362..1365  
FT /\*tag= b  
FT misc\_signal 1371..1440  
FT /\*tag= c  
FT /note= "Three copies of the binding site for  
FT transcription factors of the SGF1 forkhead type"  
FT TATA\_signal 1471..1474  
FT /\*tag= d  
FT exon 1503..1576  
FT /\*tag= e  
FT /number= 1  
FT /note= "Fibrohexamerin exon 1, encodes the signal  
FT peptide"  
FT intron 1577..2077  
FT /\*tag= f  
FT FT /number= 1  
FT /note= "Fibrohexamerin intron 1"  
FT CDS 2092..2772  
FT /\*tag= g  
FT /product= "Reporter protein Ds-Red"  
FT  
XX FR2852325-A1.  
XX PD 17-SEP-2004.  
XX PF 13-MAR-2003; 2003FR-00003137.  
XX PR 13-MAR-2003; 2003FR-00003137.  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX PA (UVLY-) UNIV LYON 1 BERNARD CLAUDE.  
XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX  
PI Chavancy G, Couble P, Durand B, Grenier A, Horard B, Julien E,  
PI Mauchamp B, Nony P, Prudhomme JC, Royer C;  
XX  
DR WPI; 2004-670997/66.  
XX  
XX New nucleic acid that directs expression of protein in silk glands of  
PT silkworm, useful for preparation of pharmaceutical proteins and for  
PT modifying textile properties of silk.  
XX  
PS Example 3; SEQ ID NO 3; 50pp; French.  
XX  
CC The present invention relates to a novel nucleic acid (I) that directs



CC expression of a protein of interest (II) specifically in the cells of the  
CC posterior silk-producing glands (A) of Bombyx mori. (I) comprises, in the  
CC 5' to 3' direction: a regulatory region (RR) comprising signals for  
CC regulating expression of a polynucleotide of interest (III) in the  
CC specified cells and, under control of RR, a region that encodes a  
CC modified signal peptide (mSP). (I) is at least 90% identical with the  
CC 1150-2026 nucleotide (nt) region of ADS75466, in which the codon  
CC containing nt 1486-1488 encodes Ala, Ile or Leu. (I) is optionally  
CC modified by insertion of 1-4 copies of the sequence ADS75467 (nucleotides  
CC 1379-1390 of ADS75466), which is the binding site for transcription  
CC factors of the SGF1 forkhead type, functioning as activator of  
CC transcription and regulator of tissue-specific expression. (III) encodes  
CC a polypeptide, specifically spidroin (of spiders) or fibroin (from  
CC organisms of the genus Galleria). (I), also related expression cassettes  
CC and vectors, are used to prepare transgenic Bombyx that secrete (II)  
CC incorporated into silk threads, e.g. where (II) is of biomedical interest  
CC (e.g. hormones, antigens, enzymes, growth factors and receptors) or where  
CC the threads are intended for use as textiles and (II) modifies resistance  
CC to ductility or elasticity. The present sequence is a modified version of  
CC ADS75466 which comprises three copies of the SGF1 forkhead type  
CC transcription factor binding site.

XX  
XX Sequence 2772 BP; 760 A; 640 C; 609 G; 763 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.24e-139 length: 2772
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x ADS75468 (1-2772)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
DB ::::
2095 GTGGCTCCTCCAAGAACGTCAAGAGATTGCGCTTCAAGTGGCATGAGGGGC 2154
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB ::::
2155 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC 2214
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrrAspIle 60
DB ::::
2215 CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTGGGACATC 2274
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
DB ::::
2275 CTGTCCCCCAGTTCAGTACGGCTTCCAAGGTGATGAGCAACCCGCCGACATCCC 2334
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrrPGLuArgValMetAsnPheGlu 100
DB ::::
2335 GACTACAGAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATCGAATTGAG 2394
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB ::::
2395 GACGGCGCGGTGGTGAACCGTGAACCAAGCACTCCTCCCTGACGACGGCTGCTCATCTAC 2454
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
DB ::::
2455 AAGGTGAAGTTTCATCGGCGTGAATTTCCCTCCGACGGCCCCCGTAAATGCAAGAAAGACC 2514
QY 141 MetGlyTrrPGLuAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
DB ::::
2515 ATGGCTGGAGAGCGCTCCACCGAGCGCTGTACCCCGCGGACGGCGTGTGAAGGGCGAG 2574
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuThr 180
DB ::::
2575 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCACTACCTGTGGAGTTCAAGTCCATC 2634
QY 181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
DB ::::
2635 TACATGGCCAGAAGACCCGCTGACGTGCCCGGCTACTACTACGTGACTCCAAGCTGAC 2694

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB ::::
2695 ATCACTCCACACAAGAGACTACACCATCGTGGAGCATACGAGCGCACCGAGGCCGC 2754
QY 221 HisHisLeuPheLeu 225
DB ::::
2755 CACCACCTGTTCTCG 2769

RESULT 31
AAL47954
ID AAL47954 standard; DNA; 4692 BP.
XX
AC AAL47954;
XX
DT 26-SEP-2002 (first entry)
XX
DE Modified yeast strain related vector pDBRedd1-N1.
XX
KW yeast; RAD54; promoter; genotoxicity cassette; cytotoxicity cassette;
KW modified yeast strain; environmental pollution; vector; ds.
XX
OS Synthetic.
XX
PN DE10061872-A1.
XX
PD 20-JUN-2002.
XX
PF 12-DEC-2000; 2000DE-01061872.
XX
PR 12-DEC-2000; 2000DE-01061872.
XX
PA (LICH/) LICHTENBERG-FRATE H.
XX
PI Lichtenberg-Frate H;
XX
DR WPI; 2002-539633/58.
XX

PT Modified yeast strain, useful for detecting toxic compounds in
PT environment, contains integrated cassettes responsive to genotoxic and
PT cytotoxic compounds.
XX

PS Example 1; Page 24-25; 34pp; German.
XX

XX The present invention relates to a modified yeast strain that contains,
XX integrated stably and functionally in its genome, a genotoxicity cassette
CC and a cytotoxicity cassette, each comprising a promoter and reporter
CC gene, both of which are different in the two cassettes. The modified
CC yeast strain is used to detect environmental pollution, especially
CC genotoxic and/or cytotoxic substances in complex environmental
CC contaminants, especially organic compounds, but also (non-)ionising
CC radiation and chemical carcinogens. Particular applications are in
CC monitoring (waste) water (e.g. as an early warning system), medical
CC toxicology screening and for industrial process control. The present
CC sequence is a vector suitable for use in the present invention
XX
SQ Sequence 4692 BP; 1111 A; 1268 C; 1247 G; 1066 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.1e-138 length: 4692
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AAL47954 (1-4692)

QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
DB ::::
682 GTGGCTCCTCCAAGAACGTCAAGAGTTTCATGCGCTTCAAGTGGCATGAGGGC 741
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40



Db 742 ACCGTGAACGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC 801  
Qy 41 HisAnthrVallylsleuLysValThrlyseGlyGlyProleuProPheAlaTrpAspIle 60  
Db 802 CACAACACCGGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTCGCTGGACATC 861  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 862 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGGCACATCCCC 921  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 922 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCCGAG 981  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 982 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGCAGAGACGGCTGCTCATCTAC 1041  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 1042 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAGACC 1101  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1102 ATGGGCTGGAGGCTCTCCACCGAGCGCTGTACCCCGCGAGCGGCTGTGAAGGGCGAG 1161  
Qy 161 IleHisLysAlaLeuLysLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1162 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCACCTACTGTGAGTTCAGTCCATC 1221  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1222 TACATGGCCAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGACTCGAAGCTGGAC 1281  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1282 ATCACTCCCAACAACGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 1341  
Qy 221 HisHisLeuPheLeu 225  
Db 1342 CACCACCTGTTCTTG 1356  
RESULT 32  
ACC44640  
ID ACC44640 standard; DNA; 4692 BP.  
XX AC ACC44640;  
XX 29-MAY-2003 (first entry)  
DT Vector pDRedin1 nucleotide sequence SEQ ID NO:29.  
XX  
DE  
XX Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; Aces; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; gene; ds.  
XX  
XX Discovered by.  
OS Synthetic.  
XX  
PN WO200297059-A2.  
XX  
XX 05-DEC-2002.  
PD  
XX  
XX 30-MAY-2002; 2002WO-US017452.  
PF  
XX  
XX 30-MAY-2001; 2001US-0294758P.  
PR 21-MAR-2002; 2002US-0366891P.  
XX  
PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
XX Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E,  
PI Stewart S, Shellard J;

XX DR WPI; 2003-140461/13.  
XX  
PT Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest.  
XX  
PS Example 1; Page 182-184; 272pp; English.  
XX  
XX The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (Aces) (II) comprising several sites that participate  
CC in recombinase catalysed recombination; and (2) a method (M1) for  
CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (I) can be used in gene therapy. (M1) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an Aces. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection,  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of Aces comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 4692 BP; 1111 A; 1268 C; 1247 G; 1066 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.1e-138 Length: 4692  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: Gaps: 0  
US-10-006-922A-12 (1-225) x ACC44640 (1-4692)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db ::::: 682 GTGCGCTCTCCAGAAAGTCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 741  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 742 ACCGTGAACGCCACGAGTTCGAGATCGAGGCGAGGGCGGCCCTTACGAGGGC 801  
Qy 41 HisAnthrVallylsleuLysValThrlyseGlyGlyProleuProPheAlaTrpAspIle 60  
Db 802 CACAACACCGGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTCGCTGGACATC 861  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 862 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGGCACATCCCC 921  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 922 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCCGAG 981  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 982 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGCAGAGACGGCTGCTCATCTAC 1041  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 1042 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAGACC 1101  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1102 ATGGGCTGGAGGCTCTCCACCGAGCGCTGTACCCCGCGAGCGGCTGTGAAGGGCGAG 1161



QY 161 ILeHISLysAlAlLeuLYsLeuLYsAspGlyGlyHisTyrLeuValGluPheLYsSerIle 180  
DB 1162 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCACCTACTGCTGGAGTTCAAGTCCATC 1221  
QY 181 TyrMetAlAlLYsLYsProValGlnLeuProGlyTyrTyrTyrValAspSerLYsLeuAsp 200  
DB 1222 TACATGGCCAAAGAACCCCGTCAGCTGCCGCTACTACTAGTGAAGCTCAAGCTGGAC 1281  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 1282 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAGTACGACGCGACCGAGGGCCGC 1341  
QY 221 HisHisLeuPheLeu 225  
DB 1342 CACCACTGTTCCTG 1356  
RESULT 33  
AADI0003  
ID AADI0003 standard; DNA; 5436 BP.  
XX AADI0003;  
AC  
XX 12-SEP-2001 (first entry)  
DT  
XX Plasmid pSFM6 to express goosecoid cDNA fused to red fluorescent protein.  
DE  
XX Mouse; fertility; reproduction; gametogenesis; microinjection; CMV;  
KW cytomagalovirus; SV40; simian virus 40; infection; goosecoid gene;  
KW plasmid pSFM 6; embryogenesis; ds.  
XX  
XX Mus musculus.  
OS Cytomegalovirus.  
OS Rhesus macaque polyoma virus.  
OS Unidentified.  
OS Chimeric.  
XX  
PN WO200148224-A1.  
XX  
PD 05-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-AU001596.  
XX  
PR 24-DEC-1999; 99AU-00004884.  
XX  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
XX  
PI Thresher R, Hinde L, Hardy C, Whyard S, Vignarajan S, Grewe PM;  
PI Patil J;  
XX  
DR WPI; 2001-425672/45.  
XX  
PT Novel construct for preventing embryogenesis in animals comprises native  
PT promoter, blocking DNA which abrogates function of crucial gene and  
PT genetic switch to regulate expression/repression of blocker/gene  
PT knockout.  
XX  
PS Example 13; Page 215-217; 241pp; English.  
XX  
XX The invention relates to a construct which allows animals to be bred in  
CC captivity but renders them infertile in the wild by allowing reversible  
CC control over fertility and reproduction. The construct comprises a native  
CC promoter, a blocking DNA sequence contoured for and designed to abrogate  
CC a crucial gene's function or to cause its mis-expression, and a genetic  
CC switch to regulate controlled expression/repression of the blocker/gene  
CC knockout. The construct is useful for preventing embryogenesis or  
CC gametogenesis in animals by stably transforming an animal cell with the  
CC construct by microinjection, transfection or infection, where the  
CC construct stably integrates into the genome by homologous recombination,  
CC and implanting the cell into a host organism, where a whole animal  
CC develops from the implanted cell. The present sequence is plasmid pSFM 6  
CC containing mouse goosecoid promoter expressing goosecoid cDNA fused to  
CC red fluorescent protein. The plasmid contains CMV promoter and SV40 polyA

CC and SV40 ori  
XX Sequence 5436 BP; 1247 A; 1525 C; 1492 G; 1172 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 1.35e-138 Length: 5436  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 4 Gaps: 0  
US-10-006-922A-12 (1-225) x AADI0003 (1-5436)  
QY 1 MetArgSerSerLYsAsnValIleLYsGluPheMetArgPheLYsValArgMetGluGly 20  
DB 1426 GTCCGCTCTCCAAGACGTTCATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 1485  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 1486 ACCGTGAACGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGCG 1545  
QY 41 HisAsnThrValLYsLeuLYsValThrLYsGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 1546 CACAACACCGTGGAAGCTGAAGTGACCAAGGGCGGCCCTGCCCTTCGCTGGACATC 1605  
QY 61 LeuSerProGlnPheGlnTyrGlySerLYsValTyrValLYsHisProAlaAspIlePro 80  
DB 1606 CTGTCCCCCAGTTCACAGTACGGCTCCCAAGGTGTACGTGAAGACACCCCGCAGCATCCCC 1665  
QY 81 AspTyrLYsLYsLeuSerPheProGluGlyPheLYsTyrPgluArgValMetAsnPheGlu 100  
DB 1666 GACTACAAAGAGCTGTCTTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 1725  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 1726 GACGGCGCGGTGTGACCTGACCCCAAGACTCTCTCTGCAAGACGGCTGCTCATCTAC 1785  
QY 121 LysValLYsPheIleGlyValAsnPheProSerAspGlyProValMetGlnLYsLYsThr 140  
DB 1786 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCGTAATGCAAGAAAGACC 1845  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLYsGlyGlu 160  
DB 1846 ATGGGCTGGAGGCGCTCCACGAGCGCGCTGTACCCCGCAGCGCGTGTGAAGGGCGAG 1905  
QY 161 ILeHISLysAlAlLeuLYsLeuLYsAspGlyGlyHisTyrLeuValGluPheLYsSerIle 180  
DB 1906 ATCCACAAGGCCCTGAAGCTGAAGAGCGGGCCACTACTGCTGAGTTCAAGTCCATC 1965  
QY 181 TyrMetAlAlLYsLYsProValGlnLeuProGlyTyrTyrTyrValAspSerLYsLeuAsp 200  
DB 1966 TACATGGCCAAAGAACCCCGTCAGCTGCCGCTACTACTAGTGAAGCTCAAGCTGGAC 2025  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 2026 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 2085  
QY 221 HisHisLeuPheLeu 225  
DB 2086 CACCACTGTTCCTG 2100  
RESULT 34  
ADE24111  
ID ADE24111 standard; DNA; 6893 BP.  
XX  
XX ADE24111;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX Proviral plasmid DNA SFr.  
DE  
XX site-specific DNA recombination; Cre recombinase;



KW	protein-transduction domain; gene activation; gene inactivation;
KW	chromosomal translocation; reporter protein; ds; circular.
OS	Synthetic.
XX	
FH	Key
FT	m1bc_RNA
FT	Location/Qualifiers
FT	1..435
FT	/tag= a
FT	/note= "plasmid pUC fragment"
FT	436..996
FT	repeat_region
FT	/tag= b
FT	/standard_name= "LTR"
FT	/note= "MPSV 5'-LTR (Delta-31)"
FT	997..1014
FT	primer_bind
FT	/tag= c
FT	/note= "Unidentified primer"
FT	1015..1560
FT	5' UTR
FT	/tag= d
FT	1561..1599
FT	m1bc_RNA
FT	/tag= e
FT	/note= "loxP1 sequence"
FT	1600..2289
FT	CDS
FT	/tag= f
FT	/product= "DeRed1"
FT	/note= "from DiScosoma sp"
FT	2290..2343
FT	m1bc_RNA
FT	/tag= g
FT	/note= "loxP2 sequence"
FT	2347..3150
FT	CDS
FT	/tag= h
FT	/product= "eGFP"
FT	/note= "from Aequorea victoria"
FT	3164..4019
FT	CDS
FT	/tag= i
FT	/product= "wPRE"
FT	4020..4080
FT	3' UTR
FT	/tag= j
FT	4081..4641
FT	repeat_region
FT	/tag= k
FT	/standard_name= "LTR"
FT	/note= "3'-LTR"
FT	4642..6893
FT	m1bc_RNA
FT	/tag= l
FT	/note= "plasmid pUC fragment"
XX	
PX	WO2003070931-A2.
PN	
XX	
PD	28-AUG-2003.
XX	
PF	19-FEB-2003; 2003WO-EP001680.
XX	
PX	21-FEB-2002; 2002DE-01007313.
PR	16-JUL-2002; 2002DE-01032196.
XX	
PA	(VISI-) VISION 7 GMBH.
XX	
PI	Baum C, Will E, Ostertag W, Klump H, Schiedlmeier B;
PX	
DR	WPI; 2003-767353/72.
XX	
PT	In vitro or in vivo site-specific DNA recombination, useful e.g. for gene
PT	inactivation, using Cre recombinase that lacks heterologous protein-
PT	transduction domain.
XX	
PS	Claim 23; SEQ ID NO 8; 84pp; German.
XX	
CC	This invention describes a novel method of site-specific DNA
CC	recombination in eukaryotic cells in vitro comprising using a
CC	bacteriophage P1 Cre recombinase that lacks heterologous protein-
CC	transduction domains. The invention also contains a reporter system for
CC	detecting site-specific DNA recombination in eukaryotic cells.
CC	Recombination with Cre recombinase lacking heterologous protein-
CC	transduction domains is used for therapeutic site-specific recombination

in eukaryotic cells, *in vivo* or *in vitro*, e.g. for partial inactivation or activation of genes, also to create a chromosomal translocation. The modified recombinase can enter intact (not electrically or chemically treated) cells when added to culture medium and catalyzes recombination events in the nucleus with recombination rate over 50%, after a single application. The Cre-recombinase does not need to be supplied from nucleic acid introduced into the cell, so its effect is immediate but short-lived, limiting cytotoxic effects; eliminating the risk of non-specific integration of DNA into the genome and providing quick results. This sequence represents the proviral plasmid Cre reporter construct SFR used in the method described in the invention.

Sequence 6893 BP; 1564 A; 1940 C; 1778 G; 1611 T; 0 U; 0 Other;

### Alignment Scores:

Pred. No.:	1.89e-138	Length:	6893
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922A-12 (1-225) x ADE24111 (1-6893)

QY 1 MetArgSerSerLysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20

Db 1603 GTGGGCTCCTCAAGACGTCATCAAGAGTTCATGCCCTCAAGGTGCCGATGGAGGC 1662

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 1663 ACCGTGAACGCCACGAGTTTCAGATCGAGGCCGAGGCCGAGGCCCTACGAGGC 1722

QY 41 HISSASnThrVallybLeuLysValThrlYsglyGlyProLeuProPheAlaTrpAspIle 60  
 Db 1723 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGCTGGAGACATC 1782

QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
 Db 1783 CTGTCTCCCCCAGTTCAGTAGGGCTCCAAGGTGATCGTGAAGCAACCCCGCCGACATCCCC 1842

Cy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
 Db 1843 GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 1902

QY 101 ASPGIGLYVALTHRVALTHRGLNASPSSERLEUGLNASPGLYCYSPHEILETYR 120

NH 1903 GACGACGGCGGTGTGACCCGTGACCTCCTCCGTGCAGACGGCTGCTCATCTAC 1962

**Qy**

121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140

|||||

1963 AACCTAGATTATCGCCGTGAATTCCCCCTCGACGGCCCCCGTAAATGCAGAAGAACACC 2022

141 MetcIyTrpGIuAlaSerThrcIuArgIeuTyPrOArgaSpGIyValIeuIysGIyGIu 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666

161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisIstyrLeuValGlnPheLysSerIle 180

181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200

201 IleThrSerHisAsnGluAspIleThrIleValGluGlnIleTyrGluArgThrGluGluArg 220  
 |||||  
 202 MetSerMetGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 225  
 |||||

Qy	221	HishisIeuPhelu	225
2203			2203

## RESULT 35



```

AAD09979/c
ID AAD09979 standard; DNA; 7910 BP.
XX
AC AAD09979;
XX
DT 12-SEP-2001 (first entry)
XX
DE pBIT(dHSP)-RFP-oHoxDS/BH construct for transfection of pacific oysters.
XX
KW Simian virus 40; SV40; fertility; reproduction; gametogenesis; fruitfly;
microinjection; infection; plasmid pBIT(dHSP)-RFP-oHoxDS/BH; oyster; HSP;
embryogenesis; RFP; heat shock promoter; red fluorescent protein; ds.
XX
OS Crassostrea gigas.
OS Rhesus macaque polyoma virus.
OS Drosophila melanogaster.
OS Unidentified.
OS Chimeric.
XX
PN WO200148224-A1.
XX
PD 05-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-AU001596.
XX
PR 24-DEC-1999; 99AU-00004884.
XX
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
PI Thresher R, Hinds L, Hardy C, Whyard S, Vignarajan S, Grewe PM;
PI Partl J;
XX
DR WPI; 2001-425672/45.
XX
PT Novel construct for preventing embryogenesis in animals comprises native
PT promoter, blocking DNA which abrogates function of crucial gene and
PT genetic switch to regulate expression/repression of blocker/gene
PT knockout.
XX
PS Example 8; Page 193-196; 241pp; English.
XX
XX The invention relates to a construct which allows animals to be bred in
XX captivity but renders them infertile in the wild by allowing reversible
XX control over fertility and reproduction. The construct comprises a native
XX promoter, a blocking DNA sequence contoured for and designed to abrogate
XX a crucial gene's function or to cause its mis-expression, and a genetic
XX switch to regulate controlled expression/repression of the blocker/gene
XX knockout. The construct is useful for preventing embryogenesis or
XX gametogenesis in animals by stably transforming an animal cell with the
XX construct by microinjection, transfection or infection, where the
XX construct stably integrates into the genome by homologous recombination,
XX and implanting the cell into a host organism, where a whole animal
XX develops from the implanted cell. The present DNA sequence is plasmid
XX construct pBIT(dHSP)-RFP-oHoxDS/BH used for transfection of Pacific
XX oysters. The plasmid comprises oyster Hox gene under the under the
XX control of Drosophila heat shock promoter (dHSP), red fluorescent protein
XX and SV40 PolyA
XX
SQ Sequence 7910 BP; 2016 A; 1893 C; 1924 G; 2077 T; 0 U; 0 Other;

Alignment Scores:
Prid. No.: 2.3e-138 Length: 7910
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: Gaps: 4

US-10-006-922A-12 (1-225) x AAD09979 (1-7910)

QY 1 MetArgSerSerLyAsnValIleIleLyGluPheMetArgPheLyValArgMetGluGly 20
:::|||||
2705 GTGCGCTCCTCCAAAGACGTCAATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGGAGGGC 2646

```

QY	21	ThrValAsnGlyHnIegIuPheGluIlegIuGlyGluGlyArgProTyrGluGly	40
Db	2645	ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC	2586
QY	41	HisAnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	2585	CACAACAACCGTGAAGCTGAAGTGACCAAGGGCGGCCCTGCCCTTGCGCTGGACATC	2526
QY	61	IeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	2525	CTGTCCCCCAGTTTCCAGTACGGCTCCAAAGGTGTAAGTGAAGCAACCCCGGACATCCCC	2466
QY	81	AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	2465	GACTACAAGAAAGCTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGTGTGAATTGAG	2406
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	2405	GACGGGGCGTGTGTAACCGTGAACCAAGACTCTCTCCCTGCAGAGCGGTGCTCATCTAC	2346
QY	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr	140
Db	2345	AAGTGAAGTTCATCGGCGGTGAATCTCCCTCGACGGCCCCGTAAATGCAAGAAAGACC	2286
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuIysGlyIu	160
Db	2285	ATGGGCTGGAGGCGCTCCACCGAGCGCTGTATCCCCCGCAGCGCGCTGTGAAGGGCGAG	2226
QY	161	IleHisIysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle	180
Db	2225	ATCCACAAGGCCCTGAAGCTGAAGAGCGGGCCCACTACTGTGTGAGTTCAAGTCCATC	2166
QY	181	TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp	200
Db	2165	TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGTACTACTACGTGACTCCAAAGCTGGAC	2106
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	2105	ATCACCCTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCAAGGGCGCGC	2046
QY	221	HisHisIeuPheIeu 225	
Db	2045	CACCACCTGTTCCTG 2031	
RESULT 36			
ABS56664	standard; DNA; 9320 BP.		
ID	ABS56664		
XX	ABS56664;		
AC			
XX	24-JAN-2003 (first entry)		
DT			
XX			
DE	Plasmid p5-Puro-CMV-(N'-EGFP)-CMV-Red(EGFP-Eu) DNA.		
XX			
KW	Vector; genotoxicity; food component; cosmetic; medicine; cancer;		
KW	genetic stability; hygromycin; ds.		
XX			
OS	Synthetic.		
XX			
PH			
Key	Location/Qualifiers		
FT	misc_feature	1..1592	
FT		/*tag= a	
FT		/note= "retroviral vector p5NM"	
FT	CDS	1617..2216	
FT		/*tag= b	
FT		/product= "puromycin resistance gene from pRetron"	
FT	Promoter	2267..2848	
FT		/*tag= c	
FT		/note= "CMV promoter from pEGFP-N1"	
FT	CDS	2906..3348	
FT		/*tag= d	
FT		/product= "N'-EGFP"	



FT	misc_feature	3374. .3392	
FT	/tag= e		
FT	/note= "retroviral vector p5NM"		
FT	3411. .3392		
FT	promoter		
FT	/tag= f		
FT	/note= "CMV promoter from pDsRed1-N1"		
FT	4038. .4718		
FT	/tag= g		
FT	/product= "pDsRed1-N1"		
FT	4766. .5508		
FT	/tag= h		
FT	/product= "EGFP-EJ"		
FT	5527. .9320		
FT	/tag= i		
FT	/note= "retroviral vector p5NM"		
XX			
PN	WO200270740-A2.		
XX			
PD	12-SEP-2002.		
XX			
PF	28-FEB-2002; 2002WO-EP002194.		
XX			
PR	05-MAR-2001; 2001DE-01010449.		
XX			
PA	(WIES/) WIESMUELLER L.		
XX			
PI	Wiesmuller L;		
XX			
DR	WPI; 2002-682909/73.		
XX			
PT	New vector containing recombination sequences, useful e.g. for assessing		
PT	compounds for genotoxicity, cancer susceptibility and mutational		
PT	frequency.		
XX			
PS	Disclosure; Page 104-106; 106pp; German.		
XX			
CC	This invention describes a novel vector containing, at least, two		
CC	homologous DNA sequences that, through DNA exchange, induce a change in a		
CC	gene present in the vector. The products of the invention are used to		
CC	assess genotoxicity of compounds (e.g. components of foods, cosmetics and		
CC	medicines), to determine susceptibility for development or progression of		
CC	cancer, for analysis of (non-)conservative recombination processes, and		
CC	joining and inactivating/reverting mutations, for characterization of		
CC	selected genes with respect to maintenance or reduction of genetic		
CC	stability, and to determine genetic (in)stability and recombinational or		
CC	mutational frequency of cells, tissues or organisms. The vector allows an		
CC	analysis to be performed in hours (making it suitable for large-scale,		
CC	routine use), and can be used with a variety of eukaryotic cells and		
CC	animals. All known types of DNA recombinations can be examined. This		
CC	sequence represents a plasmid p5-Puro-CMV-(N'-EGFP)-CMV-Red(EGFP-EJ) DNA,		
CC	described in the disclosure of the invention		
XX			
XX			
SS	Sequence 9320 BP; 2124 A; 2687 C; 2474 G; 2035 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	2.89e-138	Length:	9320
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0
US-10-006-922A-12 (1-225) x	ABSS56664 (1-9320)		
OY	1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLySValArgMetGluGly		20
	:::		
Db	4041 GTGGCGCTCCTCCAAAGACGTATCAAGAGATTGCGCTTCAAGGTGCGCATGAGGGC		4100
OY	21 ThrValAsnGlyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly		40
Db	4101 ACCGTGAACGGCCACGAGTTGAGATCGAAGGCGAAGGGCGCCCTACGAGGGC		4160
OY	41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTyrAspIle		60

Db	4161	CACAAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCCCTTCGCGTGGGACATC	4220
Qy	61	LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaSerIlePro	80
Db	4221	CTGTCCCCCAGTTCCAGTACGAGCTCCAAAGGTGTACGTGAAGCAACCCCGCAGATCCCC	4280
Qy	81	AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	4281	GACTACAAGAAGCTGTCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCAG	4340
Qy	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	4341	GACGGCGCGGTGTGACCGTGACCCAGGACTCTCTCCCTGCAGAGCGGCTGCTCATCTAC	4400
Qy	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr	140
Db	4401	AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC	4460
Qy	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIu	160
Db	4461	ATGGGCTGGAGGGCTCCACCGAGCGCCTGTACCTCCCGCGACGGCGGTGTAAGGGCGAG	4520
Qy	161	IleHisIysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle	180
Db	4521	ATCCACAAGGCCCTGAAGACTGAAGAGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC	4580
Qy	181	TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysIysAsp	200
Db	4581	TACATGGCCAAAGAGCCCGTGCAGCTGCCCCGGCTACTACTACGTGACTCCAAAGCTGGAC	4640
Qy	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIArg	220
Db	4641	ATCACTCCCAACAAGAGACTACACCATCGTGAGCAAGTACGAGCGACCGAGGGCGCG	4700
Qy	221	HisHisIysLeuPheIu 225	
Db	4701	CACCACTGTCTCTG 4715	
RESULT 37			
AD104104	ID	AD104104 standard; DNA; 9658 BP.	
XX	AD104104;		
AC			
XX	22-APR-2004	(first entry)	
DT			
DE			
XX	Vector pTnMod (Oval/Red) - Quail.		
XX			
KW	ss; vector; transposase; promoter; insertion sequence; Kozak; Tn10;		
KW	ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;		
KW	signal sequence; transgenic animal; Huntington's disease; breast cancer;		
KW	alpha-1-antitrypsin deficiency; Alzheimer's disease; congenital hypothyroidism;		
KW	cystic fibrosis; galactosaemia; congenital hypothyroidism;		
KW	maple syrup urine disease; neurofibromatosis; phenylketonuria;		
KW	sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease;		
KW	shipping fever; cattle; mastitis; bacterial; viral; infection; Types I;		
KW	type; II; diabetes.		
OS	Synthetic.		
XX			
XX			
FT	Key	Location/Qualifiers	
FT	misc_feature	1..130	
FT		/*tag= a	
FT		/note= "Remainder of F1(-) on from pBluescript11 sk(-)"	
FT	misc_feature	131..132	
FT		/*tag= b	
FT		/note= "Residue from ligation of restriction enzyme sites	
FT	promoter	used in construction"	
FT		133..1777	
FT		/*tag= c	
FT		/note= "CMV promoter/enhancer from pGwiz"	
FT	misc_feature	1778..1779	



FT	/*tag= d	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	1780. .2987	
FT	/*tag= e	/product= "Transposase"
FT		/note= "Modified from Tn10 by optimising codons for stability"
FT	2988. .2993	
FT	/*tag= f	
FT	/note= "Two engineered stop codons"	
FT	2995. .3410	
FT	/*tag= g	/note= "Taken from pGWiz"
FT	3415. .3718	
FT	/*tag= h	/note= "Non-coding DNA residual from pNK2859"
FT	3719. .3761	
FT	/*tag= i	/note= "Non-coding lambda DNA residual from pNK2859"
FT	3762. .3831	
FT	/*tag= j	/note= "Left insertion sequence recognised by Tn10"
FT	3832. .3837	
FT	/*tag= k	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	3838. .4044	
FT	/*tag= l	/note= "Multiple cloning site from pBluescript11 sk(-)"
FT	4045. .4049	
FT	/*tag= m	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	4050. .4934	
FT	/*tag= n	/note= "Japanese quail ovalbumin promoter"
FT		/note= "Includes steroid-dependant response element"
FT	4935. .4942	
FT	/*tag= o	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	4943. .5092	
FT	/*tag= p	/note= "Japanese quail ovalbumin signal sequence"
FT	5093. .5098	
FT	/*tag= q	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	5099. .5991	
FT	/*tag= r	/note= "D8Red reporter coding sequence including polyA from pD8Red1.1"
FT	5992. .5997	
FT	/*tag= s	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	5998. .6036	
FT	/*tag= t	/note= "Multiple cloning site from ZeroBlunt Topo cloning vector"
FT	6037. .6042	
FT	/*tag= u	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	6043. .6475	
FT	/*tag= v	/note= "Multiple cloning site from pBluescript11 sk(-)"
FT	6476. .6480	
FT	/*tag= w	/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	6481. .6550	
FT	insertion seq	

Seq	Feature	Annotation
XX	misc_feature	/*tag= x /note= "Right insertion sequence recognised by Tn10" 6551. .6592 /tag= y /note= "Non-coding lambda DNA residual from pNK2859" 6593. .7457 /tag= z /note= "Non-coding DNA residual from pNK2859" 7458. .9658 /tag= aa /note= "Remainder from pBluescript11 sk(-)"
XX	misc_feature	WO2004003157-A2.
XX	misc_feature	08-JAN-2004.
XX	misc_feature	26-JUN-2003; 2003WO-US020389.
XX	misc_feature	26-JUN-2002; 2002US-0392415P.
XX	misc_feature	21-JAN-2003; 2003US-0441377P.
XX	misc_feature	21-JAN-2003; 2003US-0441381P.
XX	misc_feature	21-JAN-2003; 2003US-0441392P.
XX	misc_feature	21-JAN-2003; 2003US-0441405P.
XX	misc_feature	21-JAN-2003; 2003US-0441447P.
XX	misc_feature	21-JAN-2003; 2003US-0441502P.
XX	misc_feature	(TRAN-) TRANSGENRES LLC. (LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.
XX	misc_feature	Cooper RK, Cadd GG, Fioretti WC, Deboer KF;
XX	misc_feature	WPI; 2004-083042/08.
XX	misc_feature	New vector comprising a transposase gene operably linked to a promoter, useful in treating Alzheimer's disease, breast cancer, phenylketonuria, autoimmune diseases, bacterial or viral infections or Types I and II diabetes.
XX	misc_feature	Example 4; SEQ ID NO 4; 150pp; English.
XX	misc_feature	This sequence represents a new vector, pTnMod, comprising a transposase gene operably linked to a first promoter and one or more genes of interest operably linked to one or more additional promoters, where the one or more genes of interest and their operably linked promoters are flanked by transposase insertion sequences recognized by the transposase and the first promoter comprises a modified Kozak sequence comprising ACCATG. The first transposase is a Tn10 transposase and is modified in one to twenty of the first codons. The gene of interest is operably linked to a second promoter, which is a constitutive promoter or an inducible promoter, e.g. an ovalbumin or a vitellogenin promoter. The vector further comprises a polyA sequence, a conalbumin polyA sequence, operably linked to the transposase gene. The vector further comprises two stop codons operably linked to the transposase gene. The first or second gene of interest is operably linked to a second or third promoter, respectively. The first and a second gene of interest are operably linked to a second promoter. The vector further comprises an enhancer operably linked to the one or more genes of interest. The enhancer comprises at least a portion of an ovalbumin enhancer. The vector further comprises an egg directing sequence, ovalbumin or ovomucoid signal sequence or vitellogenin targeting sequence, operably linked to the one or more genes of interest. The vector is useful in producing transgenic animals with desired proteins or molecules. They are also useful in treating Huntington's disease, alpha-1-antitrypsin deficiency, Alzheimer's disease, breast cancer, cystic fibrosis, galactosaemia, congenital hypothyroidism, maple syrup urine disease, neurofibromatosis, phenylketonuria, sickle cell disease, Smith-Lemli-Optiz Syndrome, autoimmune diseases, shipping fever in cattle, mastitis, bacterial or viral infections or Types I and II diabetes.
XX	misc_feature	Sequence 9658 BP; 2572 A; 2397 C; 2148 G; 2541 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	3.04e-138
Length:	9658



Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	12	Gaps:	0

US-10-006-922A-12 (1-225) X ADI04104 (1-9658)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	5099	GTGCGCTCTCCAGAACGTATCATCAAGGAGTTTACGCTTCAAGGTGCGCATGGAGGGC	5158
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly	40
Db	5159	ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGGAGGGCGCCCTTACGAGGGC	5218
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	5219	CACAACACCGGTGAAGCTGAAGGTGACCAAGGGGGCCCCCTCGCTCGCTGGACATC	5278
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	5279	CTGTCCCCCAGTTCAGTACGAGCTCCAAAGGTGTACGTGAAGCACCCCGCCGACATCCC	5338
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	5339	GACTACAAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGTATGAATTCGAG	5398
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	5399	GACGGCGCGTGGTGAACGTGACCCAGACTCTCTCTGACAGACGGCTTCAATCTAC	5458
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr	140
Db	5459	AAGGTGAAGTTTCGGCGTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAGAACCC	5518
QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	5519	ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCCCGCAGCGCGTGTGAAGGGCGAG	5578
QY	161	IleHisLysValAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	5579	ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGGCACTACTGCTGTGAGTTCAAGTTCATC	5638
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	5639	TACATGGCCCAAGAACCCCGTCAAGCTGCGGCTACTACTAGTGAAGTCCAAAGCTGAC	5698
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	5699	ATCACTCCACACAGAGACTACACCATCTGTGAGCAGTACGAGCGACCGAGGGCGGC	5758
QY	221	HisHisLeuPheLeu 225	
Db	5759	CACCACCTGTCTCTG 5773	
RESULT 38			
AD104103			
ID	AD104103	standard; DNA; 9678 BP.	
XX	AC	AD104103;	
XX	DT	22-APR-2004 (first entry)	
XX	DE	Vector pTnMod (Oval/Red)-Chicken.	
XX	BS	vector; transposase; promoter; insertion sequence; Kozak; Tn10;	
KW	ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;		
KW	signal sequence; transgenic animal; Huntington's disease;		
KW	alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer;		
KW	cystic fibrosis; galactosaemia; congenital hypothyroidism;		
KW	maple syrup urine disease; neurofibromatosis; phenylketonuria;		
KW	sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease;		

KW		shipping fever; cattle; mastitis; bacterial; viral; infection; Types I;
XW		type; II; diabetes.
XX	Synthetic.	
OS		
FH	Key	Location/Qualifiers
FT	misc_feature	1..130 /*tag= a /note= "Remainder of Fl(-) on from pBluescriptl sk(-)" 131..132 /*tag= b /note= "Residue from ligation of restriction enzyme sites used in construction" 133..1777 /*tag= c /note= "CMV promoter/enhancer from pGwiz" /note= "CMV promoter is modified by the addition of an ACC sequence upstream of ATG" 1778..1779 /*tag= d /note= "Residue from ligation of restriction enzyme sites used in construction"
FT	promoter	1780..2987 /*tag= e /product= "Transposase" /note= "Modified from Tn10 by optimising codons for stability" 2988..2993 /*tag= f /note= "Two engineered stop codons" 2994 /*tag= g /note= "Residue from ligation of restriction enzyme sites used in construction"
FT	CDS	2995..3410 /*tag= h /note= "Taken from pGWiz" 3415..3718 /*tag= i /note= "Non-coding DNA residual from pNK2859" 3719..3761 /*tag= j /note= "Non-coding lambda DNA residual from pNK2859"
FT	polyA_signal	3762..3831 /*tag= k /note= "Left insertion sequence recognised by TnlO"
FT	misc_feature	3832..3837 /*tag= l /note= "Residue from ligation of restriction enzyme sites used in construction"
FT	misc_feature	3838..4044 /*tag= m /note= "Multiple cloning site from pBluescriptll sk(-)" /note= "May be used to insert the sequence of interest into the vector"
FT	misc_feature	4045..4049 /*tag= n /note= "Residue from ligation of restriction enzyme sites used in construction"
FT	misc_signal	4050..4951 /*tag= o /note= "Upstream elements"
FT	misc_signal	4960..5112 /*tag= p /note= "Chicken ovalbumin signal sequence"
FT	misc_feature	5113..5118 /*tag= q /note= "Residue from ligation of restriction enzyme sites used in construction"
FT	CDS	5119..6011 /*tag= r /note= "Dared reporter coding sequence including polyA from pDsRed1.l"
FT		







DB	5779	CACCACCTGTTCTCTG	5793
RESULT	39		
ID	AD104102	standard; DNA; 10263 BP.	
AC	AD104102;		
DT	22-APR-2004	(first entry)	
DE	Vector pTnMod (CMV/Red).		
XX	ss; vector; transposase; promoter; insertion sequence; Kozak; Tn10;		
XX	ovalbumin; vitellogenin; polyA; conalbumin; enhancer; ovomucoid;		
XX	signal sequence; transgenic animal; Huntington's disease;		
XX	alpha-1-antitrypsin deficiency; Alzheimer's disease; breast cancer;		
XX	cyatic fibrosis; galactosaemia; congenital hypothyroidism;		
XX	maple syrup urine disease; neurofibromatosis; phenylketonuria;		
XX	sickle cell disease; Smith-Lemli-Opitz Syndrome; autoimmune disease;		
XX	shipping fever; cattle; mastitis; bacterial; viral; infection; Types I;		
XX	type; II; diabetes.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	misc_feature	1..130	
FT		/tag= a	
FT		/note= "Remainder of Fl(-) on from pBluescript11 SK(-)"	
FT	misc_feature	131..132	
FT		/tag= b	
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"	
FT	promoter	133..177	
FT		/tag= c	
FT		/note= "CMV promoter/enhancer from pGwiz"	
FT		/note= "CMV promoter is modified by the addition of an ACC sequence upstream of ATG"	
FT	misc_feature	1778..1779	
FT		/tag= d	
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"	
FT	CDS	1780..2987	
FT		/tag= e	
FT		/product= "Transposase"	
FT		/note= "Modified from Tn10 by optimising codons for stability"	
FT	misc_feature	2988..2993	
FT		/tag= f	
FT		/note= "Two engineered stop codons"	
FT	misc_feature	2994	
FT		/tag= g	
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"	
FT	polyA_signal	2995..3410	
FT		/tag= h	
FT		/note= "Taken from pGwiz"	
FT	misc_feature	3415..3718	
FT		/tag= i	
FT		/note= "Non-coding DNA residual from pNK2859"	
FT	misc_feature	3719..3761	
FT		/tag= j	
FT		/note= "Non-coding lambda DNA residual from pNK2859"	
FT	insertion_seq	3762..3831	
FT		/tag= k	
FT		/note= "left insertion sequence recognised by Tn10"	
FT	misc_feature	3832..3837	
FT		/tag= l	
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"	
FT	misc_feature	3838..4044	
FT		/tag= m	
FT		/note= "Multiple cloning site from pBluescript11 SK(-)"	

FT		/note= "May be used to insert the sequence of interest into the vector"
FT	promoter	4049. .5693
FT		/*tag= n
FT		/note= "CMV promoter/enhancer from pGwiz"
FT	misc_feature	5694. .5701
FT		/*tag= o
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	CDS	5702. .6617
FT		/*tag= p
FT		/note= "DsRed reporter coding sequence including polyA from pDsRed1.1"
FT	misc_feature	6618. .7101
FT		/*tag= q
FT		/note= "Multiple cloning site from pBluescript11 sk(-)"
FT		/note= "May be used to insert the sequence of interest into the vector"
FT		7102. .7106
FT	misc_feature	/*tag= r
FT		/note= "Residue from ligation of restriction enzyme sites used in construction"
FT	insertion_seq	7107. .7176
FT		/*tag= s
FT		/note= "Right insertion sequence recognised by Tn10"
FT	misc_feature	7177. .7218
FT		/*tag= t
FT		/note= "Non-coding lambda DNA residual from pNK2859"
FT	misc_feature	7219. .8062
FT		/*tag= u
FT		/note= "Non-coding DNA residual from pNK2859"
FT	misc_feature	8063. .10263
FT		/*tag= v
FT		/note= "Remainder from pBluescript11 sk(-)"
XX	WO2004003157-A2.	
XX		
XX	08-JAN-2004.	
PD		
XX		
XX	26-JUN-2003; 2003WO-US020389.	
XX		
PR	26-JUN-2002; 2002US-0392415P.	
PR	21-JAN-2003; 2003US-0441377P.	
PR	21-JAN-2003; 2003US-0441381P.	
PR	21-JAN-2003; 2003US-0441392P.	
PR	21-JAN-2003; 2003US-0441405P.	
PR	21-JAN-2003; 2003US-0441447P.	
PR	21-JAN-2003; 2003US-0441502P.	
XX		
PA	(TRAN-) TRANSGENES LLC.	
PA	(LOU) UNIV LOUISIANA STATE & AGRIC & MECH COLL.	
XX		
PI	Cooper RK, Cadd GG, Fioretti WC, Deboer KF;	
XX		
DR	WPI; 2004-083042/08.	
XX		
PT	New vector comprising a transposase gene operably linked to a promoter, useful in treating Alzheimer's disease, breast cancer, phenylketonuria, autoimmune diseases, bacterial or viral infections or Types I and II diabetes.	
XX		
PS	Example 2; SEQ ID NO 2; 150pp; English.	
XX		
CC	This sequence represents a new vector, pTnMod (CMV/Red), comprising a transposase gene operably linked to a first promoter and one or more genes of interest operably linked to one or more additional promoters, where the one or more genes of interest and their operably linked promoters are flanked by transposase insertion sequences recognized by the transposase and the first promoter comprises a modified Kozak sequence comprising ACCATG. The transposase is a Tn10 transposase and is modified in one to second of the first codons. The gene of interest is operably linked to a second promoter, which is a constitutive promoter or an inducible promoter, e.g. an ovalbumin or a vitellogenin promoter. The vector	



CC further comprises a polyA sequence, a conalbumin polyA sequence, operably  
CC linked to the transposase gene. The vector further comprises two stop  
CC codons operably linked to the transposase gene. The first or second gene  
CC of interest is operably linked to a second or third promoter,  
CC respectively. The first and a second gene of interest are operably linked  
CC to a second promoter. The vector further comprises an enhancer operably  
CC linked to the one or more genes of interest. The enhancer comprises at  
CC least a portion of an ovalbumin enhancer. The vector further comprises an  
CC egg directing sequence, ovalbumin or ovomucoid signal sequence or  
CC vitellogenin targeting sequence, operably linked to the one or more genes  
CC of interest. The vector is useful in producing transgenic animals with  
CC desired proteins or molecules. They are also useful in treating  
CC Huntington's disease, alpha-1-antitrypsin deficiency, Alzheimer's  
CC disease, breast cancer, cystic fibrosis, galactosaemia, congenital  
CC hypothyroidism, maple syrup urine disease, neurofibromatosis,  
CC phenylketonuria, sickle cell disease, Smith-Lemli-Optiz Syndrome,  
CC autoimmune diseases, shipping fever in cattle, mastitis, bacterial or  
CC viral infections or Types I and II diabetes.

SQ Sequence 10263 BP; 2648 A; 2595 C; 2374 G; 2646 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.31e-138 Length: 10263  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 12 Gaps: 0

US-10-006-922A-12 (1-225) x ADI04102 (1-10263)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20

Db 5725 GTGCGCTCTCCAGAAGCGTCATCAAGAGTTTCATGCGCTTCAAGTCGCGCATGAGGGC 5784

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

Db 5785 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 5844

QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60

Db 5845 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCTTGGCTGGACATC 5904

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80

Db 5905 CTGTCCCCCAGTTCACGATACGGCTCCCAAGTGTAAGAACACCCCGCACATCCCC 5964

QY 81 AspTyrTrpValLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100

Db 5965 GACTACAGAAGCTGTCTTCCCGAGGCTTCAAGTGGAGCGCGGTGATGAATTGAG 6024

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

Db 6025 GACGGCGGCTGTGTGACCGTGACCCAGGACTCTCCCTGCAGGACGGCTTCTCATCTAC 6084

QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140

Db 6085 AAGGTGAAGTTTCATCGGCGTGAACTTCCCTCCGACGGCCCGTAAATGCAGAAGAGACC 6144

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlu 160

Db 6145 ATGGGCTGGGAGGCTTCACCGAGCGCTGTACCCCGCGAGCGGTGCTGAAGGGCGAG 6204

QY 161 IleHisLeuAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180

Db 6205 ATCCACAAGGCCCTGAAGCTGAAGGACGGGGCCACTACCTGGTGAGTTCAAGTCCATC 6264

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 6265 TACATGGCCAAGAACCCCGTGACGCTGCCCGCTACTACTACGTGACTCCAAGCTGGAC 6324

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 6325 ATCACCCTCCACAACGAGGAGTACACCATCTGTGAGCAGTACGAGCGCACCGAGGCCGC 6384

QY 221 HisHisLeuPheLeu 225

Db 6385 CACCACTGTTCCTG 6399

RESULT 40

ABA95906

ID ABA95906 standard; cDNA; 681 BP.

AC ABA95906;

DT 29-MAY-2002 (first entry)

DE Yeast optimised RFP encoding cDNA SEQ ID NO 2.

XX Yeast; red fluorescent protein; RFP; plant; transgenic; GFP;

KW Saccharomyces cerevisiae; Nicotiana tabacum; Arabidopsis thaliana;

KW Escherichia coli; green fluorescent protein; biotechnology; gene; ss.

XX Anthozoa.

OS

XX DE20001395-U1.

XX 15-MAR-2001.

XX 27-JAN-2000; 2000DE-02001395.

XX 27-JAN-2000; 2000DE-02001395.

XX (GPCB-) GPC BIOTECH AG.

XX WPI; 2002-228394/29.

XX P-PSDB; ABB08835.

XX New DNA encoding red fluorescent protein, useful as marker in

XX biotechnology, has sequence optimized for expression in eukaryotes,

XX especially yeast or plants.

XX Claim 2; Fig 1; 19pp; German.

XX The invention relates to DNA (I) containing either sequence ABA95905 or

XX sequence ABA95906 encoding a yeast optimised Red Fluorescent Protein

XX (YRFP). (I) are used to express red fluorescent protein (RFP) in

XX eukaryotes, especially yeast, especially Saccharomyces cerevisiae and

XX plants, especially dicotyledonous plants including Nicotiana tabacum or

XX Arabidopsis thaliana and also in prokaryotes, especially bacteria,

XX especially Escherichia coli. RFP is useful in the same way as green

XX fluorescent protein but is more generally applicable in modern

XX biotechnology. (I) are optimised for expression in yeast and so generate

XX RFP at higher levels with stronger fluorescence and thus lowers the

XX detection limit and gives a better signal-to-noise ratio

SQ Sequence 681 BP; 198 A; 147 C; 161 G; 175 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.65e-140 Length: 681  
Score: 1209.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.59% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABA95906 (1-681)

QY 2 ArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGlyThr 21

Db 7 AGATCTTCTAAGAAGCGTCATCAAGAAATTTCATGAGATTCAAGTTAGAATGAAGTACT 66







|||||  
Db 607 ACCUCUACAACGAGACUACUACUGCGAACAAUACGACGUAAGGUAAGACAC 666  
QY 222 Hisleupheleu 225  
|||||  
Db 667 CACUGUCUUCUG 678  
RESULT 42  
AAD28208  
ID AAD28208 standard; cDNA; 678 BP.  
XX  
AC AAD28208;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. fluorescent protein E5 encoding cDNA.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; E5; S8.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "E5 protein"  
XX  
PN WO200196373-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US019097.  
XX  
PR 14-JUN-2000; 2000US-0211607P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Fradkov AF, Terzikikh A;  
XX  
DR WPI, 2002-154595/20.  
DR P-PSDB; AAE17541.  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
PS Disclosure; Fig 2; 89pp; English.  
XX  
CC The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene  
CC expression during development of a multicellular organism or during  
CC cellular differentiation, in response to a drug or other inducer of  
CC promoter activity, as a reporter to serve as a read-out of promoter  
CC activity, monitoring intracellular protein movement or translocation,  
CC protein trafficking, or protein stability, to investigate temporal  
CC aspects of the activity of a regulatory element, for determining cell  
CC fate during development and organ remodeling, in spatial and temporal  
CC visualisation of newly synthesised proteins and accumulated proteins, and  
CC in distinguishing between newly formed and pre-existing structures, e.g.  
CC membrane junctions and extracellular matrix components. The fluorescent  
CC timer proteins may further be used to investigations where photobleaching  
CC techniques are employed, as detectable labels, as selectable markers, as  
CC biosensors in prokaryotic and eukaryotic cells, in protease cleavage  
CC assays, and as second messenger detectors. The nucleic acids can be used  
CC to generate transgenic, non-human plants or animals or site-specific gene  
CC modifications in cell lines. The present sequence is a cDNA encoding

CC Discosoma sp. E5 fluorescent protein derived from humanised wild-type  
CC Anthozoa protein dFFP583 by substituting Val to Ala at 105 and Ser to Thr  
CC at 197  
XX  
SQ Sequence 678 BP; 147 A; 224 C; 203 G; 104 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.7e-139 Length: 678  
Score: 1207.00 Matches: 223  
Percent Similarity: 99.56% Conservative: 1  
Best Local Similarity: 99.11% Mismatches: 1  
Query Match: 99.42% Indels: 0  
DB: Gaps: 0  
US-10-006-922A-12 (1-225) x AAD28208 (1-678)  
QY 1 MetArgSerSerLySAAsnValIleLySGluPhMetArgPheLySValArgMetGluGly 20  
Db 1 ATGCGCTCCTCCAAGAACGTATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGCG 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 121 CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGGCCCTGCGCTTGGCGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAAGTTCAGTAGGGCTCCAAGGTGTACTGAAGCACCCCCGCATATCCCC 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySThrGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGGCGTGGCGACCGGTGACCCAGACTCTCCCTCGACGAGCGGCTGCTCATCTAC 360  
QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTGAAGTTCATCGGCGGTGAATCTCCCTCCGACGGCCCGGTGATGAGAAGAAGACC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 421 ATGGGCTGGGAGGCGCTCCACCGAGCGGCTGTACCCCGGCAAGCGGCTGTGAAGGCGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
Db 481 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCTACTACTGTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGCTACTACTAAGTGACACCAAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACAAGAGACTACCATCATGTGAGCAGTACGAGCGCACCGAGGGCGCG 660  
QY 221 HisHisleupheleu 225  
Db 661 CACCACTGTTCCTG 675  
RESULT 43  
AAH47656  
ID AAH47656 standard; cDNA; 681 BP.  
XX  
AC AAH47656;  
XX  
DT 30-NOV-2001 (first entry)  
XX  
DE Anthozoa red fluorescent protein synthetic mutant nucleotide sequence.



XX Fluorescent protein; Anthozoan; fluorescence; marker; FRET; mutant; ss.  
XX Synthetic.  
XX WO200162919-A1.  
XX  
XX 30-AUG-2001.  
XX  
XX 13-FEB-2001, 2001WO-US004625.  
XX  
XX 23-FEB-2000, 2000US-0184732P.  
XX  
XX (AURO-) AURORA BIOSCIENCES CORP.  
XX  
XX Nelson D, Zamatra E, Tsien R;  
XX  
XX MPI; 2001-557704/62.  
XX  
XX P-PSDB; AAG65510.  
XX  
XX  
XX Proteins for Fluorescence Resonance Energy Transfer (FRET) comprise  
XX functional red fluorescent proteins, and the encoding nucleic acids, with  
XX key mutations for improving the proteins function.  
XX  
XX  
XX Disclosure; Page 86; 90pp; English.  
XX  
XX The invention provides a nucleic acid encoding functional red fluorescent  
XX CC protein (II) that differs from the sequence of an Anthozoan red  
XX fluorescent protein by at least one amino acid substitution, and with  
XX different fluorescent properties. The red fluorescent protein of the  
XX invention can be expressed by standard recombinant methodology. (II) are  
XX used a fluorescent markers and FRET partners. It is used for identifying  
XX protein-protein interactions. (II) is also suitable for multiplexed  
XX fluorescent analysis and FRET-based applications using existing Aequorea  
XX fluorescent proteins. (II) has improved brightness, reduced spectral  
XX cross talk, and is rapidly and efficiently expressed in mammalian cells.  
XX The key mutations in the encoding nucleic acids provide improved folding,  
XX brightness, and create (II) with sharper, more defined excitation and  
XX emission peaks when expressed in mammalian cells. The present sequence  
XX represents the nucleotide sequence of an improved synthetic mutant of an  
XX anthozoan fluorescent protein  
XX  
XX Sequence 681 BP; 160 A; 201 C; 220 G; 100 T; 0 U; 0 Other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 3.02e-139 Length: 681  
XX Score: 1205.00 Matches: 223  
XX Percent Similarity: 99.56% Conservative: 1  
XX Best Local Similarity: 99.11% Mismatch: 1  
XX Query Match: 99.26% Indels: 0  
XX DB: 4 Gaps: 0  
XX  
XX US-10-006-922A-12 (1-225) x AAH47656 (1-681)  
XX  
QY 1 MetArgSerSerLyAsnValIleLySGluPhMetArgPheLySValArgMetGluGly 20  
XX :::  
Db 4 GTGAGGAGCAGCAAGACGTGATCAAGAGTTCAAGGTGCGCATGGAAGGC 63  
XX  
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGluGly 40  
XX :::  
Db 64 ACCGTGAACGGCCACGAGCTTCGAGATCGAGGCGAGGGCGAGGCCCTCAGAGGGC 123  
XX  
QY 41 HisAsnThrVallybLeuLySValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60  
XX :::  
Db 124 CACAAACACCGTGAAAGCTTAAGGTGACCAAGGGCGGCCCTCGCCTGGGACATC 183  
XX  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrVallybHisProAlaAspIlePro 80  
XX :::  
Db 184 CTGAGCCCCCATTTCAGTACGCGCAGCAAGGTGTACGTGAAGCAACCCCGCGACATCCCC 243  
XX  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTrpGluArgValMetAsnPheGlu 100  
XX :::  
Db 244 GACTACAAAGAAAGCTGAGCTTCCCGAGGGCTTCAAGTGGAAGAGGGTGATGAACCTTCAG 303

QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
XX :::  
Db 304 GACGGCGCGTGTGACCGTGACCAGACAGACAGACCGCTTGACGAGCGCTTCATCTAC 363  
XX  
QY 121 LysVallybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140  
XX :::  
Db 364 AAGGTAAAGTTCATCGCGGTGAACCTCCCGACGACGCGCCCGTGATGCAGAAAGAAC 423  
XX  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleLybGlyGlu 160  
XX :::  
Db 424 ATGGGCTGGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGCGAG 483  
XX  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
XX :::  
Db 484 ATCCACAAAGCCCTGAAGCTGAAGAGACGCGCGCACTACTGTGAGTTCAAGTCCATC 543  
XX  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
XX :::  
Db 544 TACATGGCCAAAGACCCGTGCACTGCGCTGCTACTACTACTGGAAGTCCAAAGCTGGAC 603  
XX  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
XX :::  
Db 604 ATCACCAAGCCACAAAGAGACTACCATCGTGAGCAGTACGAGAGAGACCGAGGCGCAG 663  
XX  
QY 221 HisHisLeuPheLeu 225  
XX :::  
Db 664 CACCACCTGTCTCTG 678  
XX  
XX  
XX RESULT 44  
XX ADR30814  
XX ID ADR30814 standard; DNA; 7508 BP.  
XX  
XX AC ADR30814;  
XX  
XX AC ADR30814;  
XX  
XX DT 04-NOV-2004 (first entry)  
XX  
XX DE Zebrafish BDNF gene construct DNA #2.  
XX  
XX KW Brain-derived neurotrophic factor; BDNF; promoter; zebrafish; ds.  
XX  
XX OS Danio rerio.  
XX  
XX OS Chimeric.  
XX  
XX OS Unidentified.  
XX  
XX  
XX Key  
FH misc\_feature 1..20 Location/Qualifiers  
FT misc\_feature 1..20  
FT FT /\*tag= a  
FT /\*note= "Vector DNA"  
FT misc\_feature 15..5104  
FT FT /\*tag= b  
FT /\*note= "Expression vector DNA"  
FT misc\_feature 21..1776  
FT FT /\*tag= c  
FT /\*note= "Zebrafish BDNF gene 5' flank region"  
FT promoter 1777..1804  
FT FT /\*tag= d  
FT /\*note= "Zebrafish BDNF promoter gene"  
FT exon 1805..2099  
FT FT /\*tag= e  
FT /\*note= "Exon 1c (5' UT) of Zebrafish BDNF gene "  
FT misc\_feature 2100..2119  
FT FT /\*tag= f  
FT /\*note= "Vector DNA"  
FT misc\_feature 2120..2815  
FT FT /\*tag= g  
FT /\*note= "Reporter vector DNA"  
FT misc\_feature 2816..2820  
FT FT /\*tag= h  
FT /\*note= "Linker DNA"  
FT misc\_feature 2821..5099  
FT FT /\*tag= i  
FT /\*note= "3' flank region"



FT misc\_feature 5100..7508  
FT /tag= j  
FT /note= "Vector DNA"  
XX US2004157294-A1.  
PN XX  
PD 12-AUG-2004.  
XX  
PF 23-DEC-2003; 2003US-00742828.  
XX  
PR 30-NOV-2001; 2001US-0334079P.  
PR 29-NOV-2002; 2002US-00306737.  
XX  
PA (USGO ) US DEPT VETERANS AFFAIRS.  
XX  
PI Heinrich G, Huynh G;  
XX  
DR WPI; 2004-592772/57.  
XX  
PT Nucleic acid construct useful for screening substances for their ability  
PT to influence production of brain derived neurotrophic factor (BDNF),  
PT comprises promoter of BDNF gene and red fluorescent reporter.  
XX  
XX  
PS Claim 1; SEQ ID NO 4; 56bp; English.  
XX  
CC The present invention provides an isolated brain-derived neurotrophic  
CC factor (BDNF) gene promoter and its associated nucleic acid constructs.  
CC The invention is useful for screening a number of substances for their  
CC ability to influence the production of BDNF in living organisms. The  
CC present sequence is zebrafish brain-derived neurotrophic factor (BDNF)  
CC gene construct DNA. This sequence is used in the invention.  
XX  
SQ Sequence 7508 BP; 2249 A; 1604 C; 1665 G; 1990 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.57e-137 Length: 7508  
Score: 1203.00 Matches: 222  
Percent Similarity: 99.56% Conservative: 2  
Best Local Similarity: 98.67% Mismatches: 1  
Query Match: 99.09% Indels: 0  
DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x ADR30814 (1-7508)

Qy 1 MetArgSerSerLyAsnValIlelyeGluphMetArgphelyValArgMetGlucly 20  
Db 2123 GTGCGCTCTCCAAAGACGTATCAAGAGCTTATGCGCTTCAAGTCCGCATGAGGGC 2182  
Qy 21 ThrValaenglyHisGluphegluilegluglygluglyArgProTyrGlucly 40  
Db 2183 ACCGTGAAGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 2242  
Qy 41 HisAsnThrVallybLeuLyAsnValThrlyeGlyGlyProleuProphelaTTPaspile 60  
Db 2243 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCTTGGCTGGACATC 2302  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLySerValTyrVallyHisProAlaAspIlePro 80  
Db 2303 CTGTCCCCCGAGTTCACGATACGAGTGTGTAAGACACCCCGACATCCCC 2362  
Qy 81 AspTyrlyLeuLySerPheProGluGlyPhelySTrpGluArgValMetAsnPhelGlu 100  
Db 2363 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 2422  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerleuGlnAspGlyCysPheileTyr 120  
Db 2423 GACGGCGGGCGTGGCGACCGTGACCCGAGACTTCTCCCTGCAAGACGGCTGCTTCACTAC 2482  
Qy 121 LySVallybPheileGlyValAsnPheproSeraSpGlyProValMetGlnLyblySThr 140  
Db 2483 AAGGTGAAGTTATCGGGCGTGAACCTTCCCTTCGACGCGCCCGTGATGCAAGAAAGACC 2542  
Qy 141 MetGlyTTrpGluAlaSerThrGluArgleuTyrProArgAspGlyValleuLyblyGlu 160

Db 2543 ATGGGCTGGAGGCGCTCCACCGAGCGCGCTGTACCCCGCGACGCGGTGTGAAGGGGAG 2602  
Qy 161 IleHisLyAlaLeuLybLeuLyAspGlyGlyHisTyrleuValGluPhelySerile 180  
Db 2603 ATCCACAAGGCCCTGAAGCTGAAGGACGCGCGGCGCACTACTGTGTGAGTTCAAGTCAATC 2662  
Qy 181 TyrMetAlaLybLybProValGlnleuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
Db 2663 TACATGGCCAAAGAGCCGTGACGCTGCCCGGCTACTACTAGTGGACACCAAGCTGGAC 2722  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 2723 ATCACTCCCAACAAGAGACTACACCATCGTGGAGAGTACGAGCGCACCGAGGGCGC 2782  
Qy 221 HisHisLeuPheLeu 225  
Db 2783 CACCACTGTCTCTG 2797

RESULT 45  
ACA55359  
ID ACA55359 standard; DNA; 6985 BP.

AC ACA55359;  
XX  
DT 06-JUN-2003 (first entry)  
XX  
XX Transformation vector piggyBAC related plasmid p3xP3-DsRed-orf.  
DE  
XX piggyBac; transposon; eukaryotic transformation vector; ds;  
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
KW circular.  
XX  
OS Synthetic.

XX  
XX US2002173634-A1.  
XX  
XX 21-NOV-2002.  
PD  
XX 30-OCT-2001; 2001US-00001189.  
PF  
XX 31-OCT-2000; 2000US-0244677P.  
PR 01-NOV-2000; 2000US-0244984P.  
PR  
XX  
XX

PA (FRAS/) FRASER M J.  
PA (LIXX/) LI X.  
PA (BEAM/) BEAM T.  
PA (HUA/) HUA-VAN A.

PI Fraser MJ, Li X, Beam T, Hua-Van A;  
XX  
XX WPI; 2003-352597/33.  
DR

XX  
PT New DNA molecule in the transposon piggyBac, useful for transferring  
PT genes into host cells or embryos for transforming the cells of embryos  
PT that can be used in making transgenic organisms.

XX Example 10; Fig 9(B); 151pp; English.

XX  
CC The invention describes a DNA molecule comprising at least 163  
CC consecutive nucleotide base pairs of the 3' terminal region beginning at  
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base  
CC pairs of the 5' terminal region beginning at the 5'terminal base pair, of  
CC the piggyBac molecule. The region extends from the restriction site SacI  
CC to the end of the piggyBac molecule. The DNA molecule in the transposon  
CC piggyBac is useful for transferring genes into host cells or embryos for  
CC transforming the cells of embryos. The transformed cells or embryos are  
CC useful for developing or making transgenic organisms. This sequence  
CC represents plasmid used in the creation of minimal sequence eukaryotic  
CC transformation vector piggyBac

SQ Sequence 6985 BP; 1860 A; 1680 C; 1731 G; 1714 T; 0 U; 0 Other;



Alignment Scores:

Pred. No.: 7.86e-137 Length: 6985  
Score: 1197.00 Matches: 224  
Percent Similarity: 99.56% Conservative: 1  
Best Local Similarity: 99.12% Mismatches: 0  
Query Match: 98.60% Indels: 1  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ACA55359 (1-6985)

QY 1 MetArgSerSerLyAsnValIleLyEGluPheMetArgPheLyValArgMetGluGly 20  
DB 2974 GTGCGCTCTCCAAAGAACGTCAATCAAGAGATTATGCGCTTCAAGGTGGCATGAGAGGC 3033  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40  
DB 3034 ACCGTGAAGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGCCCTTACGAGGGC 3093  
QY 41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 3094 CACAACACCGTGAAAGCTGAAGGTGACCAAGGCGGCCCTTGCTTGGCTGGACATC 3153  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80  
DB 3154 CTGTCCCCCGAGTCCAGTACGGCTCCAAAGGTGTAAGTGAAGCACCCGCGACATCCCC 3213  
QY 81 AspTyrLyLeuLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 3214 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 3273  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
DB 3274 GACGGCGCGGTGTGACCGTGACCCAGGACTCTCTCTGCAAGACGGCTGCTTCACTAC 3333  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
DB 3334 AAGGTGAAGTTCATCGGCGTGAACTTCCCTCCGACGGCCCGTAATGCAAGAAAGACC 3393  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 3394 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 3453  
QY 161 IleHisLyS-AlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
DB 3454 ATCCACAAGGGCCCTGAAGCTGAAGAGCGCGGCCCACTACCTGTGAGTTCAGTCCAT 3513  
QY 180 eTyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAs 200  
DB 3514 CTACATGGCCCAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGACTCCAAAGCTGGA 3573  
QY 200 PileThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyAr 220  
DB 3574 CATCACCTCCCAAGAGAGACTACACCATCGTGAGCAGTACGAGCGCACGAGGGCGG 3633  
QY 220 GHISHisLeuPheLeu 225  
DB 3634 CCACCACTGTCTCTG 3649

RESULT 46

ABL61142  
ID ABL61142 standard; DNA; 666 BP.  
XX

AC ABL61142;  
XX

DT 18-SEP-2002 (first entry)  
XX

DE Red fluorescent protein reporter DNA fragment.  
XX

KW Alpha-SMA; smooth muscle actin; screening; smooth muscle cell;  
KW myofibroblast; gene therapy; red fluorescent protein; ds.  
XX

OS Synthetic.  
XX

PN EP1172375-A1.  
XX  
PD 16-JAN-2002.  
XX  
PF 22-DEC-2000; 2000EP-00128446.  
XX  
PR 11-JUL-2000; 2000DE-01033633.  
PR 31-OCT-2000; 2000DE-01053879.  
XX  
PA (ODEN/) ODONTAL M.  
XX  
PI Odenthal M, Jung D;  
XX  
DR WPI; 2002-149590/20.  
XX

PT New nucleic acid containing regulatory region of the smooth muscle actin  
PT gene, useful e.g. for manipulating gene expression in smooth muscle  
PT cells.

PS Disclosure; Page 20; 44p; German.

CC This invention describes a novel nucleic acid (I) comprising: (i) at  
CC least one functional region (Ia) from the regulatory region of the alpha-  
CC smooth muscle actin (SMA) gene and (ii) at least one additional  
CC functional sequence (Ib) operably linked to (Ia). The products of the  
CC invention can be used for preparing genetically modified eukaryotic cells  
CC or organisms, for isolation and screening of smooth muscle cells,  
CC myofibroblasts or related cells, and for manipulation of gene expression  
CC and/or cell function in smooth muscle cell or myofibroblasts,  
CC particularly for gene therapy. Component (Ia) provides cell-type- or  
CC differentiation-specific expression or modulation of genes. This sequence  
CC represents a DNA fragment of red fluorescent protein (EGFP) which can be  
CC used as a reporter molecule under the control of the alpha-SMA described  
CC in the invention

XX SQ Sequence 666 BP; 202 A; 126 C; 176 G; 162 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.82e-138 Length: 666  
Score: 1196.00 Matches: 221  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.52% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x ABL61142 (1-666)

QY 5 LySAsnValIleLySGluPheMetArgPheLySValArgMetGluGlyThrValAsnGly 24  
DB 1 AAGAATGTTATCAAGAGATTCAATGAGTTTAAAGTTTCGATGGAAGCAACGTCATATGGG 60  
QY 25 HisGluPheGluIleGluGlyGlyGlyArgProTyrGlyGlyHisAsnThrVal 44  
DB 61 CACGAGTTTGAATAAGAGCGAAGAGAGAGGAGCCATACGAAGGCCAATAACCGTA 120  
QY 45 LySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGln 64  
DB 121 AAGCTTAAGGTAAACCAAGGGGAGACCTTGCCATTGCTTGGGATATTGTCACACCAA 180  
QY 65 PheGlnTyrGlySerLySValTyrValLyHisProAlaAspIleProAspTyrLySLeu 84  
DB 181 TTTCAGTATGGAAGCAAGTATATGTCAAGCACCCCTGCCACATACGACGATATAAAG 240  
QY 85 LeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGluAspGlyGlyVal 104  
DB 241 CTGTCAATTTCTGAAGATTAAATGGGAAGGTCATGAACCTTGAAGACGGTGGCGTC 300  
QY 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyrLySValLySPhe 124  
DB 301 GTTACTGTAAACCAAGATTCAGTTTGACGATGGCTGTTTCATCTCAAGGTCAAGTTC 360  
QY 125 IleGlyValAsnPheProSerAspGlyProValMetGlnLySThrMetGlyTyrGlu 144



Db 361 ATGGCGTGAAC TTCTCCGATGGA CCTGTTATGCAAAAGAGA CAATGGCTGGGA 420  
Qy 145 Alse rThrglu argLeu TyrPro argAsp GlyValLeuLysGlyGluIleHisLysAla 164  
Db 421 GCCAGCACTGAGCGTTGTATCCCTCGATGATGGCGTGTGAAAGAGAGATTCA TAAGGCT 480  
Qy 165 LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPhelySerIleTyrMetAlaLys 184  
Db 481 CTGAAGCTGAAAGACGGGTGTCATTACTAGTTGAATTCAAAAGTATTACATGCGCAAG 540  
Qy 185 LysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHis 204  
Db 541 AAGCTGTGCAGCTACCAAGGTA CTATATGTTGACTCCAAACTGGATATTAACAAGCCAC 600  
Qy 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHisLeuPhe 224  
Db 601 AACGAACTATACAAATCGTTGAGCAGTATGAAAGAACCGAGGACGCCACCATCTGTTC 660  
Qy 225 Leu 225  
Db 661 CTT 663  
RESULT 47  
ADO78067  
ID ADO78067 standard; DNA; 711 BP.  
XX  
AC ADO78067;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Corallimorpharia red fluorescent protein mutant Redi DNA.  
XX  
KW fluorescent protein; red fluorescent protein; green fluorescent protein;  
KW gene expression detection; regeneration; multiple labeling system;  
KW in vivo marker; microinjection assay; taxonomic marker; colour indicator;  
KW food additive; cosmetic; Actinodiscus; Discosoma; gene; ds; mutant; Redi;  
KW coral-like anemone.  
XX  
OS Corallimorpharia.  
XX  
FH Key location/Qualifiers  
FT CDS 1..711  
FT /\*tag= a  
FT /product= "Corallimorpharia red fluorescent protein  
FT mutant Redi"  
XX  
PN US2004110225-A1.  
XX  
PD 10-JUN-2004.  
XX  
PF 09-DEC-2002; 2002US-00314936.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
XX  
PI Gtbbs PDL, Carter RW, Schmale MC;  
XX  
DR MPI; 2004-467659/44.  
DR P-PSDB; ADO78068.  
XX  
PT New first mutant fluorescent protein having an optical property  
PT relatively different to a corresponding optical property produced by a  
PT wild-type red or green fluorescent protein, useful for detecting  
PT expression of a gene.  
XX  
PS Claim 9; SEQ ID NO 1; 30pp; English.  
XX  
CC The invention describes a first mutant fluorescent protein (I) comprising  
CC an amino acid sequence that differs from that of a corresponding wild-  
CC type red fluorescent protein by an amino acid substitution, where the

CC first mutation fluorescent protein mutant has an optical property that  
CC differs relative to the corresponding optical property produced by the  
CC wild-type red or green fluorescent protein. (I) is useful for detecting  
CC expression of gene. The polynucleotide (II) encoding (I) is useful for  
CC detecting expression of a gene which involves introducing (II) into a  
CC cell or organism, allowing the cell to replicate, and detecting  
CC expression of the nucleic acid by emission of fluorescent light. The  
CC expression of the nucleic acid expression is detected in vivo or in  
CC vitro. (II) is useful for regenerating an animal such as a zebrafish  
CC which contains (II). (I) is useful as a marker for detecting expression  
CC of a gene, in biochemical assays, and as reagents. (I) is also useful in  
CC multiple labeling systems, as in vivo markers such as in mRNA  
CC microinjection assays, and as taxonomic markers for studies of genetics,  
CC colour indicators in diagnostic kits, coloured food additives, and  
CC cosmetics ingredients. (I) is useful in research for up or down  
CC regulation, to monitor promoter activity, to allow longer term monitoring  
CC and to localise proteins. (I) has enhanced properties such as  
CC substantially enhanced fluorescence and reduced toxicity. This sequence  
CC encodes Redi, a mutant of the red fluorescent protein isolated from an  
CC aquatic species believed to be either an Actinodiscus or Discosoma  
CC species.  
XX  
SQ Sequence 711 BP; 214 A; 131 C; 189 G; 177 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 4.18e-138 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 12 Gaps: 0  
US-10-006-922a-12 (1-225) x ADO78067 (1-711)  
Qy 1 MetargSerSerLysAanValIleLysGluPhMetArgPhelysValArgMetGluGly 20  
Db 1 ATGAGTTGTTCCAAGATGTATTATCAAGAGGTTCA TGAGGTTTAAAGTTCGTATGAAAGGA 60  
Qy 21 ThrValAsnGlyHisGluPhGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAGGCGAAGAGAGGAGCCATACGAAGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATAACGTAAGCTTAAGGTAACCAAGGGGGACCTTGGCATTTGTTGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCACAATTTCA GTATGGAAGCAAGGTATGTCAAGCATCTCGCGACATACCA 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPhelysTrpGluArgValMetAsnPhGlu 100  
Db 241 GACTATAAAAGCTGTCAITTTCTGAAGATTAAATGGGAAAGGGTCATGAAC TTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCGTACTGTAAACCCAGGATTCAGTTTGACGAGATGGCTGTTTCATCTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGTCAAGTTCATTGGCGTGAAC TTCTCTGTGATGACCTGTTATGCAAAAGAGACA 420  
Qy 141 MetGlyTrpGluAlaSerThrgluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGCTGGGAAGCCAGCACTGAGCGTTGTATCCCTCGATGATGGCGTGTGAAAGAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPhelySerIle 180  
Db 481 ATTCAATAAGGCTCTGAAGTTGAAAGACGGGTGTCATTACTAGTTGAATTCAAAACTATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAAAGAGCCTGTGCAGCTACCAAGGTA CTATATGTTGACTCCAAACTGAT 600



Oy		201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrglGlyArg	220
			::	
Dd		601	ATAACAAGCCACAAACAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGGACGC	660
Oy		221	HISHisLeuPheLeu	225
Dd		661	CACCATCTGTTCCTT	675
RESULT	48			
ID	ADO78069		standard; DNA; 711 BP.	
XX	AD078069;			
XX	DT	26-AUG-2004	(first entry)	
DE			Corallimorpharia red fluorescent protein mutant RedII DNA.	
XX				
KW			fluorescent protein; red fluorescent protein; green fluorescent protein;	
KM			gene expression detection; regeneration; multiple labeling system;	
KM			in vivo marker; microinjection assay; taxonomic marker; colour indicator;	
KM			food additive; cosmetic; Actinodiscus; Discosoma; gene; ds; mutant;	
KW			RedII; coral-like anemone.	
XX				
OS			Corallimorpharia.	
XX				
FH	Key		Location/Qualifiers	
FT	CDS		1..711	
FT			/tag= a	
FT			/product= "Corallimorpharia red fluorescent protein	
FT			mutant RedII"	
XX				
PN			US2004110225-A1.	
XX				
PD			10-JUN-2004.	
XX				
PF			09-DEC-2002; 2002US-00314936.	
XX				
PR			09-DEC-2002; 2002US-00314936.	
XX				
PA			(GIBB/) GIBBS P D L.	
PA			(CART/) CARTER R W.	
PA			(SCHM/) SCHMALE M C.	
XX				
PI			Gibbs PDL, Carter RW, Schmale MC;	
XX				
DR			WPI; 2004-467659/44.	
DR			P-PSTDB; ADO78070.	
PT				
PT			New first mutant fluorescent protein having an optical property	
PT			relatively different to a corresponding optical property produced by a	
PT			wild-type red or green fluorescent protein, useful for detecting	
PT			expression of a gene.	
XX				
PS			Claim 10; SEQ ID NO 3; 30pp; English.	
XX				
CC			The invention describes a first mutant fluorescent protein (I) comprising	
CC			an amino acid sequence that differs from that of a corresponding wild-	
CC			-type red fluorescent protein by an amino acid substitution, where the	
CC			first mutation fluorescent protein mutant has an optical property that	
CC			differs relative to the corresponding optical property produced by the	
CC			wild-type red or green fluorescent protein. (I) is useful for detecting	
CC			expression of gene. The polynucleotide (II) encoding (I) is useful for	
CC			detecting expression of a gene which involves introducing (II) into a	
CC			cell or organism, allowing the cell to replicate, and detecting	
CC			expression of the nucleic acid by emission of fluorescent light. The	
CC			expression of the nucleic acid expression is detected in vivo or in	
CC			vitro. (II) is useful for regenerating an animal such as a zebrafish,	
CC			which contains (II). (I) is useful as a marker for detecting expression	
CC			of a gene, in biochemical assays, and as reagents. (I) is also useful in	
CC			multiple labeling systems, as in vivo markers such as in mRNA	
CC			microinjection assays, and as taxonomic markers for studies of genetics,	

CC	colour indicators in diagnostic kits, coloured food additives, and cosmetics ingredients. (I) is useful in research for up or down regulation, to monitor promoter activity, to allow longer term monitoring and to localise proteins. (I) has enhanced properties such as substantially enhanced fluorescence and reduced toxicity. This sequence encodes Redit, a mutant of the red fluorescent protein isolated from an aquatic species believed to be either an <i>Actinodiscus</i> or <i>Discosoma</i> species.		
XX	Sequence 711 BP; 215 A; 131 C; 188 G; 177 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	4.18e-138	Length:	711
Score:	1196.00	Matches:	221
Percent Similarity:	99.11%	Conservative:	2
Best Local Similarity:	98.22%	Mismatches:	2
Query Match:	98.52%	Indels:	0
DB:	12	Gaps:	0
US-10-006-922A-12 (1-225) x ADO78069 (1-711)			
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20	
Db	1 ATGAGTTGTTCCAAGAATGTTATCAAGAGTTTCATGAGTTTAAGTTTGGTATGGAAGA	60	
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40	
Db	61 ACCGTCAATGGGACGAGTTTGAATAGAAGCGGAAGAGAAGGAGCCATACGAAGGC	120	
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60	
Db	121 CACAATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCCTTGCCATTGCTGGGATATT	180	
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80	
Db	181 TTGTCAACCACAATTTCAATGGAAGCAAGTATATGTCAGCAATCTGCCACATACCA	240	
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100	
Db	241 GACTATAAAAGCTGTCTATTTCTGAAGATTTAATGGAAAGGTCATGAACCTTGAA	300	
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerSerLeuGlnAspGlyCysPheIleTyr	120	
Db	301 GACGGTGGCGTCTGTAAGTAAACCAAGATTTCCAGTTTGCAAGATGGCTTTCATCTAC	360	
QY	121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140	
Db	361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCTGATGACCTGTTATGCAGAAAGAGACA	420	
QY	141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160	
Db	421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAATGGCGTTGAAGAGAG	480	
QY	161 IleHisLysAlaLeuLysLeuLysAspGlyHisTyrLeuValGluPheLysSerIle	180	
Db	481 ATTCAATAGGCTCTGAAGTTGAAGACGCGTGTCTATTACCTAGTTGAATCAAACTATT	540	
QY	181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200	
Db	541 TACATGGCAAGAAAGCCGTGTGACGCTACAGGGTACTACTATGTGACTCCAACTGGAT	600	
QY	201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220	
Db	601 ATAACAAGCCACAACAAGACTATACATCTTGAGCAGTATGAAAGAACGAGGAGCGC	660	
QY	221 HisHisLeuPheLeu 225		
Db	661 CACCATCTGTTCCCTT 675		
RESULT 49			
ADZ84222			
ID	ADZ84222	standard; DNA; 711 BP.	
XX			



AC AD284222;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE Actinodiscus/Discosoma RFP mutant Red I-encoding DNA, SEQ ID NO:1.  
XX  
KM Red fluorescent protein; Red I; mutant; gene; ds.  
XX  
OS Corallimorpharia.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..711  
FT /\*tag= a  
FT /product= "Mutant Ac/DsRFP Red I"  
FT replace(694,A)  
FT mutation /\*tag= b  
FT /note= "This nucleotide is A in the Red II mutant,  
FT resulting in Asn rather than Asp at position 232 of the  
FT protein"  
FT  
PN US2005100954-A1.  
XX  
PD 12-MAY-2005.  
XX  
PF 23-DEC-2004; 2004US-00021014.  
XX  
PR 09-DEC-2002; 2002US-00314936.  
XX  
PA (GIBB/) GIBBS P D L.  
PA (CART/) CARTER R W.  
PA (SCHM/) SCHMALE M C.  
XX  
PI Gibbs PDL, Carter RW, Schmale MC;  
XX  
XX WPI, 2005-345399/35.  
DR P-PSDB; AD284223.  
XX  
PT Novel mutant fluorescent protein comprising amino acid sequence that  
PT differs from wild-type red or green fluorescent protein by amino acid  
PT substitutions, useful as markers for detecting desired gene expression.  
XX  
PS Claim 3; SEQ ID NO 1; 29pp; English.  
XX  
XX The invention relates to mutant red and green fluorescent proteins which  
CC have higher fluorescent intensities compared to the wild-type proteins  
CC from which they are derived. The invention also discloses nucleic acids  
CC encoding the mutant fluorescent proteins, and transgenic animals  
CC comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
CC of the invention, Red I (AD284223) and Red II (AD284225), are derived  
CC from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
CC mushroom coral believed to be either an Actinodiscus or Discosoma  
CC species, while the mutant green fluorescent proteins (GFP), Green I  
CC (AD284227) and Green (AD284229), are derived from a novel wild-type GFP  
CC (known as McGFP) isolated from the great star coral Montastraea  
CC cavernosa. The Ac/DsRFP and McGFP mutants were obtained by subjecting the  
CC wild-type Ac/DsRFP and McGFP cDNAs to low-stringency PCR to introduce  
CC random mutations, cloning the PCR products into bacterial expression  
CC vectors, and selecting the transformed bacteria having higher relative  
CC fluorescence compared to those expressing the wild-type proteins. The  
CC mutant fluorescent proteins of the invention may be used in a wide range  
CC of applications. They are useful as markers for detecting the expression  
CC of a gene of interest, or can be used as in vivo markers in mRNA  
CC microinjection assays in transgenic animals. They can be used in  
CC expression systems (e.g., as components of fusion proteins), in multiple  
CC labeling systems, or can be used in biochemical assays and as reagents.  
CC They may additionally be used as taxonomic markers for studies of  
CC cnidarian genetics, as color indicators in diagnostic kits, as colored  
CC food additives and as cosmetic ingredients. The present sequence  
CC represents DNA encoding the mutant Actinodiscus/Discosoma red fluorescent  
CC protein Red I.  
XX  
XX Sequence 711 BP; 214 A; 131 C; 189 G; 177 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.18e-138 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 14 Gaps: 0  
US-10-006-922A-12 (1-225) x AD284222 (1-711)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAAGAATGTTATCAAGAGTTCATGAGGTTTAAGGTTCTGATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGGCCAAGAGAGGAGGCCATACGAAGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGACCTTTGCCATTGTGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 181 TTGTACACCACAATTTCAATGGAAGCAAGGTATTATGCAAGCATCCTGCCAGATTACA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAGCTGTCAATTCCTGAGAGATTAAATGGAAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCGTCGTACTGTAACCAAGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCATTTGGCGGTGAACCTTCCTTGATGCACTGTTATGCAAAAGACACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGCTGGGAAGCCAGACTGAGCGTTGTATCCTGATGCGCTGTGAAGGAGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 481 ATTCAATAAGGCTCTGAAGTTGAAGAAGCGGTGTCAATTACCTAATTGAATTCAAACTAAT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAAGCCTTGCAAGCTACCAAGGTACTACTATGTTGACTCCAAACTGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATTAACAAGCCACACAAAGACTATACAATCGTTGAGCAGATATGAAGAACCAGAGGAGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675  
RESULT 50  
AD284224  
ID AD284224 standard; DNA; 711 BP.  
XX  
AC AD284224;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE Actinodiscus/Discosoma RFP mutant Red II-encoding DNA, SEQ ID NO:3.  
XX  
KM Red fluorescent protein; Red II; mutant; gene; ds.  
XX  
OS Corallimorpharia.  
OS Synthetic.



XX Key Location/Qualifiers  
FH 1..711  
FT CDS /\*tag= a  
FT /product= "Mutant Ac/DsRFP Red II"  
FT replace(694,G)  
FT /\*tag= b  
FT /note= "This nucleotide is G in the Red I mutant,  
FT resulting in Asp rather than Asn at position 232 of the  
FT protein"  
XX PN US2005100954-A1.  
XX PD 12-MAY-2005.  
XX PF 23-DEC-2004; 2004US-00021014.  
XX PR 09-DEC-2002; 2002US-00314936.  
XX PA (GIBB/) GIBBS P D L.  
XX PA (CART/) CARTER R W.  
XX PA (SCHM/) SCHMALE M C.  
XX PI Gibbs PDL, Carter RW, Schmale MC;  
XX WI PFI, 2005-345399/35.  
XX DR P-PSDB, ADZ84225.  
XX PT Novel mutant fluorescent protein comprising amino acid sequence that  
PT differs from wild-type red or green fluorescent protein by amino acid  
PT substitutions, useful as markers for detecting desired gene expression.  
XX PS Example 4; SEQ ID NO 3; 29pp; English.  
XX CC The invention relates to mutant red and green fluorescent proteins which  
CC have higher fluorescent intensities compared to the wild-type proteins  
CC from which they are derived. The invention also discloses nucleic acids  
CC encoding the mutant fluorescent proteins, and transgenic animals  
CC comprising such a nucleic acid. The mutant red fluorescent proteins (RFP)  
CC of the invention, Red I (ADZ84223) and Red II (ADZ84225), are derived  
CC from a novel wild-type RFP (referred to as Ac/DsRFP) isolated from a  
CC mushroom coral believed to be either an Actinodiscus or Discosoma  
CC species, while the mutant green fluorescent proteins (GFP), Green I  
CC (ADZ84227) and Green (ADZ84229), are derived from a novel wild-type GFP  
CC (known as MCGFP) isolated from the great star coral Montastraea  
CC cavernosa. The Ac/DsRFP and MCGFP mutants were obtained by subjecting the  
CC wild-type Ac/DsRFP and MCGFP cDNAs to low-stringency PCR to introduce  
CC random mutations, cloning the PCR products into bacterial expression  
CC vectors, and selecting the transformed bacteria having higher relative  
CC fluorescence compared to those expressing the wild-type proteins. The  
CC mutant fluorescent proteins of the invention may be used in a wide range  
CC of applications. They are useful as markers for detecting the expression  
CC of a gene of interest, or can be used as in vivo markers in mRNA  
CC microinjection assays in transgenic animals. They can be used in  
CC expression systems (e.g., as components of fusion proteins), in multiple  
CC labeling systems, or can be used in biochemical assays and as reagents.  
CC They may additionally be used as taxonomic markers for studies of  
CC cnidarian genetics, as color indicators in diagnostic kits, as colored  
CC food additives and as cosmetic ingredients. The present sequence  
CC represents DNA encoding the mutant Actinodiscus/Discosoma red fluorescent  
CC protein Red II. Red II was found to have a fluorescent intensity at least  
CC 50% greater than the mutant RFP Red I (ADZ84223).  
XX SQ Sequence 711 BP, 215 A; 131 C; 188 G; 177 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 4.18e-138 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 14 Gaps: 0

US-10-006-922A-12 (1-225) x ADZ84224 (1-711)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAAGAATGTATCAAGAGAGTTCATGAGGTTTAAGGTTGGTATGAAGGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCACGAGTTTGAAATAGAAAGCGAAGAGAGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGACCCTTGCCATTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 TTGTACACCACAATTCAGTATGAGCAAGCATATATGTCAAGCATTCGCCGACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAGCTGTCTATTCCTGAAGCATTTAAATGGGAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGTGGCTGTTACTGTAAACCCAGCATTCAGTTTGCAGGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 361 AAGGTCAAGTTCAATTGGCGTGAACCTTCTTGATGACCTGTTATGCAAAAGAAAGACA 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGTGAAGGAGAG 480  
QY 161 IleHisValAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
DB 481 ATTCATAAGCTCTGAAGTTGAAGACGGGTGCATTACCTAGTTGAATTCAAACTATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCAAGAACCTGTGCAAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg 220  
DB 601 ATAACAAGCCACAACAAGACTATACATCGTTGACAGATATGAAGAACCGAGGACGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCTT 675  
RESULT 51  
AAD46282  
ID AAD46282 standard; DNA; 678 BP.  
XX AAD46282;  
AC 27-DEC-2002 (first entry)  
XX DE Discosoma sp. drFP583 (NFP-6) mutant DNA, ES-NA.  
XX KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; gene;  
KW mutant; ds.  
XX KM  
XX OS Discosoma sp.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH 1..678  
FT CDS /\*tag= a



FT /product= "drFP583 mutant protein, E5-NA"  
XX WO200268459-A2.  
PN  
XX 06-SEP-2002.  
PD  
XX  
PF 20-FEB-2002; 2002WO-US005749.  
XX  
XX  
PR 21-FEB-2001; 2001US-0270983P.  
PR 04-DEC-2001; 2001US-00006922.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov S, Lukyanov K, Yanushkevich Y, Savitsky A, Fradkov A;  
XX  
XX WPI; 2002-691654/74.  
DR P-PSDB; AAE28837.  
DR  
XX  
XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.  
XX  
XX  
PS Disclosure; Page 73; 80pp; English.  
PS  
XX  
XX The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is Discosoma sp. drFP583 (NFP-6) mutant DNA of the invention  
XX  
SQ Sequence 678 BP, 145 A; 225 C; 203 G; 105 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 1.63e-137 Length: 678  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0  
DB: 6 Gaps: 0  
  
US-10-006-922A-12 (1-225) x AAD46282 (1-678)  
  
QY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
DB 1 ATGGCCTCTCCGAGAACGTCATCACCGAGTTCATGCGTTCAGGTGCGCATGAGGGC 60  
QY 21 ThrValaAngLyHisGluPheGluIleGlyGlyGlyGlyArgProTyrGluGly 40  
DB 61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAACACCGGTGAAGTTGAAGGTGACCAAGGCGGCCCTGCTTCGCTGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCGAGTTCAGTACGAGCTCCAGGTGTACGTGAAGCACCCCGACATCCCC 240  
QY 81 AspTyrLyValySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACAAGAGCTGTCTCCCGAGGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGGGCTGGCGAACCGTGAACCCAGACTCTCCCTGCAGGAGCGGCTCTCATCTAC 360  
QY 121 LysValLySPheIleGlyValaSnPheProSerAspGlyProValMetGlnLySThr 140

DB 361 AAGGTGAAGTTCATCGGCGGTGAACCTTCCCTCCGACGAGCCCGGTGATGCAGAGAGACC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGAGGCGCTCCACCGAGCGCGCTGTACCCCGGACGCGGTGTGAAGGGCGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuAsp 180  
DB 481 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValaSPSerLySLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCGTGCAAGCTGCCCGGCTACTACTACGTGACACCAAGCTGGAC 600  
QY 201 IleThrSerHisAngLyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGAGGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACTGTCTCTG 675  
  
RESULT 52  
AAD28209  
ID AAD28209 standard; cDNA; 678 BP.  
XX  
XX AAD28209;  
AC  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Discosoma sp. non-aggregating mutant timer protein E5NA encoding cDNA.  
XX  
KW Fluorescent timer protein; protein movement; translocation; trafficking;  
KW promoter activity; gene expression; transgenic plant; gene modification;  
KW protein age; mutant; E5NA; ss.  
XX  
OS Discosoma sp.  
OS Synthetic.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "Mutant E5NA protein"  
XX  
XX WO200196373-A2.  
XX  
XX  
PD 20-DEC-2001.  
XX  
XX 13-JUN-2001; 2001WO-US019097.  
PF  
XX 14-JUN-2000; 2000US-0211607P.  
PR  
XX  
PA (CLON-) CLONTECH LAB INC.  
PI Fradkov AF, Tersikh A;  
XX  
XX WPI; 2002-154595/20.  
DR P-PSDB; AAB17542.  
DR  
XX  
XX  
PT New fluorescent timer proteins comprising an emission spectrum that  
PT changes over time from a first wavelength to a second wavelength, useful  
PT for monitoring intracellular protein movement, translocation, trafficking  
PT or stability.  
XX  
XX  
PS Disclosure; Fig 3; 89pp; English.  
PS  
XX  
XX The invention relates to a fluorescent timer protein having an emission  
CC spectrum that changes over time after synthesis from a first wavelength  
CC to a second wavelength. The fluorescent timer proteins are useful in  
CC monitoring the activity of a promoter, determining the age of a protein,  
CC identifying an agent that modulates the activity of a promoter and in  
CC enriching a population of cells comprising a fluorescent timer protein.  
CC The fluorescent timer proteins are also useful for assessing gene



expression during development of a multicellular organism or during cellular differentiation, in response to a drug or other inducer of promoter activity, as a reporter to serve as a read-out of promoter activity, monitoring intracellular protein movement or translocation, protein trafficking, or protein stability, to investigate temporal aspects of the activity of a regulatory element, for determining cell fate during development and organ remodelling, in spatial and temporal visualisation of newly synthesised proteins and accumulated proteins, and in distinguishing between newly formed and pre-existing structures, e.g. membrane junctions and extracellular matrix components. The fluorescent timer proteins may further be used to investigations where photobleaching techniques are employed, as detectable labels, as selectable markers, as biosensors in prokaryotic and eukaryotic cells, in protease cleavage assays, and as second messenger detectors. The nucleic acids can be used to generate transgenic, non-human plants or animals or site-specific gene modifications in cell lines. The present sequence is a cDNA encoding Discosoma sp. E5NA non-aggregating fluorescent timer protein derived from CC fluorescent timer protein E5 by substituting amino acids at positions CC R2A, K5E and K9T. E5 protein is derived from humanised wild-type Anthozoa CC protein drFP583 by substituting Val to Ala at 105 and Ser to Thr at 197 XX  
SQ Sequence 678 BP, 145 A; 226 C; 203 G; 104 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.63e-137 Length: 678
Score: 1191.00 Matches: 220
Percent Similarity: 98.67% Conservative: 2
Best Local Similarity: 97.78% Mismatches: 3
Query Match: 98.11% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x AAD28209 (1-678)

QY 1 MetArgSerSerHisAsnValIleLeuGluPheMetArgPheLeuValArgMetGluGly 20
DB 1 ATGGCCTCCTCCGAGAACGTGATCACCAGATTGCGCTTCAAGGTGCGCATGAGGGC 60
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGluGlyGluGlyArgProTyrGluGly 40
DB 61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGGCGGCGCCCTTACGAGGGC 120
QY 41 HisAsnThrValLeuLeuLeuValThrLeuGlyGlyProLeuProPheAlaTrpAspIle 60
DB 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCGCCCTGCTCGCTGGGACATC 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLeuValTyrValIleHisProAlaAspIlePro 80
DB 181 CTGTCCCCCAGTTCAGTAGCGGCTCCAAGGTGATGTAAGCACCCTCCGACATCCCC 240
QY 81 AspTyrLeuLeuSerPheProGluGlyPheLeuTyrGluArgValMetAsnPheGlu 100
DB 241 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATACTTCGAG 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 301 GACGGCGGCGTGGCGACCGTGAACCAAGAACTCTCCCTGCAAGGACGGCTGCTTCTAC 360
QY 121 LysValLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB 361 AAGGTGAAGTTCATCGGCGGTGAACCTTCCCTCCGACGGCGCCGTGATGCAAGAAAGACC 420
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB 421 ATGGGCTGGAGGCGCTCCACGAGCGCCTGTACCCTCCGCGACGGCGTGTGAAGGCGAG 480
QY 161 IleHisLeuAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 481 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACTACTGGTGAAGTTCAAGTTCATC 540
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB 541 TACATGGCCCAAGAGCCCGTGCAGCTGCGCGGCTACTACTACGTGAGACACCAAGCTGAGC 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 601 ATCACTCCCAACAGAGAGACTACACCATCGTGAAGCAGTAGACGACCGAGGGCCGC 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCACCTGTTCTCTG 675

RESULT 53
AAD46281
ID AAD46281 standard; DNA; 675 BP.
XX AAD46281;
AC
XX 27-DEC-2002 (first entry)
DT
XX Discosoma sp. drFP583 (NFP-6) mutant DNA, E57-NA.
DE
XX Fluorescent protein; chromoprotein; protease cleavage assay; filter;
KW fluorescence activated cell sorting application; fluorescent timer;
KW biosensor; fluorescence resonance energy transfer application; FRET;
KW colouring agent; recombinant DNA application; analyte detection assay;
KW sunscreen; second messenger detector; drFP583 protein; NFP-6; gene;
KW mutant; ds.
XX
XX Discosoma sp.
OS Synthetic.
OS
XX Key Location/Qualifiers
FH CDS 1..675
FT /\*tag= a
FT /product= "drFP583 mutant protein, E57-NA"
FT /note= "No stop codon"
FT /partial

WO200268459-A2.

06-SEP-2002.

20-FEB-2002; 2002WO-US005749.

21-FEB-2001; 2001US-0270983P.
PR 04-DEC-2001; 2001US-00006922.

(CLON-) CLONTECH LAB INC.

Lukyanov S, Lukyanov K, Yanushevich Y, Savitsky A, Fradkov A;

WPI: 2002-691654/74.

P-PSDB; AAE28836.

New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant of an aggregating Chidarian chromo- or fluorescent protein or mutant for analyte detection assays or fluorescence activated cell sorting applications.

Disclosure; Page 73; 80pp; English.

The invention relates to nucleic acid molecules encoding non-aggregating chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are useful in analyte detection assays, as colouring agents, as markers in recombinant DNA applications, as sunscreens or filters, in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in screening assays, as second messenger detectors, in fluorescence activated cell sorting applications, in protease cleavage assays or as fluorescent timers. The present sequence is Discosoma sp. drFP583 (NFP-6) mutant DNA of the invention

SQ Sequence 675 BP, 142 A; 227 C; 204 G; 102 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.74e-137 Length: 675
Score: 1186.00 Matches: 219



Percent Similarity:	98.22%	Conservative:	2
Best Local Similarity:	97.33%	Mismatches:	4
Query Match:	97.69%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AAD46281 (1-675)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGGACCTCTCCGAGAACGTATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTCAGATCGAGGCGAGGGCGCGCCCTACGAGGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACACACACCGTGAAGCTGAAGGTGACCAAGGGCGCCCTGCCCTTCCGCTGGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCAGTACGGCTCCAGGTGTACGTGAAGCACCCCGCCGACATCCCC	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTTCAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGGCGTGGCGACCGTGACCCAGAGACTCTCTCCGACGAGACGGCTGCTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	AAGGTGAAGTTCATCGGCGTGAACTTCCCTCCGACGGCCCCGTGATGCAGAGAGAAC	420
QY	141	MetGlyTyrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	ATGGGCTGGGAGGCTCCACCGAGCGCTGTACCCCCGCGACGGCGTGCTGAAGGGCGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGTGAGTTCAAGTTCATC	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCCAAGAGCCCGTGACGTGCCCGGTACTACTACGTGAGCGCCAAAGCTGGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	601	ATCACTTCCCACAAAGGAGTACACCATGTGTGAGACAGTACGAGCGCACCGAGGGCCGC	660
QY	221	HisHisLeuPheLeu 225	
Db	661	CACCACTGTCTCTG 675	

RESULT 54	
ADV96699	
ID	ADV96699 standard; DNA; 747 BP.
XX	
AC	ADV96699;
XX	
DT	10-MAR-2005 (first entry)
XX	
DE	Plant cell gene silencing related dsRed gene coding region.
XX	
KW	Gene silencing; plant; short interfering RNA; siRNA; RNA interference;
KW	phage resistance; ds; gene; dsRed.
XX	
OS	Synthetic.
XX	
PN	US2004261149-A1.
XX	
PD	23-DEC-2004.

XX	24-FEB-2004; 2004US-00785862.
PF	
XX	
PR	24-FEB-2003; 2003US-0449646P.

PA	(FAUQ/)	FAUQUET C M.
PA	(PADM/)	PADMANABHAN C.
PA	(RAMA/)	RAMACHANDRAN V.

Fauguet CM, Padmanabhan C, Ramachandran V,

WPI; 2005-038811/04.

**PT** Suppressing germiniviral genes in a plant cell comprises introducing a short interfering RNA into the plant cell.

Example 1; SEQ ID NO 10; 15pp; English.

The invention relates to a novel method for suppressing the expression of a target gene in a plant cell, comprising introducing a short interfering RNA (siRNA) into the plant cell. The method for suppressing the expression of a target gene in a plant cell comprises: introducing a siRNA into the plant cell, where the siRNA is a double-stranded molecule with a first strand consisting of a nucleic acid sequence which is substantially complementary to a nucleic acid sequence of a target gene and a second strand consisting of a nucleic acid sequence which is substantially complementary to the first strand, where the siRNA is capable of suppressing the expression of the target gene. The method, siRNA and composition are useful for suppressing expression of a target gene in a plant cell. Particularly, they are useful for suppressing or inhibiting viral gene expression, specifically germiniviral gene expression, in plant cells. In a particular embodiment of the invention, siRNAs are targeted to the coding region (AC1) of the replication-associated protein (Rep) of an ACWV-CM strain. This polynucleotide sequence represents the coding region of a gene used in the plant gene suppression method of the invention.

Sequence 747 BP; 159 A; 248 C; 224 G; 116 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	7.78e-137
Score:	1186.00
Percent Similarity:	98.22%
Best Local Similarity:	97.33%
Query Match:	97.69%
DB:	14
	Gaps:
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

US-10-006-922A-12 (1-225) x ADV96699 (1-747)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
		:::	
Db	1	ATGGCCTCTCCGAGAACGTCATCACCGATTGCGCTTCAAGTGCGCATGGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGGCCACGAGTTGCGATCGAGGGCGAGGGCGGGCCCTTACGAGGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	121	CACACACCCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCGCTGGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTGTCCCCCAGTTCACGTAAGGCTCCAAAGTGTAAGTGAAGCACCCCGCCGACATCCCC	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	241	GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACTTGAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGCGGTGGCGACCGGTGACCAAGGACTCTCCCTGCAGAGACGGCTGTCATCTAC	360



QY

121

LyValAllybPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr

140

DB

361

AAGTGAAGTTTCATCGCGGTGAAGTTCCCTCCGACGGCCCCGTGATGCAAGAAGACACC

420

QY

141

MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

DB

421

ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGTGTGAAGGGCGAG

480

QY

161

IleHisLysAlaLeuLysLeuLysLeuAspGlyLysHisTyrLeuValGluPheLysSerIle

180

DB

481

ACCCACAAGGCCCTGAAGCTGAAGAGCGGGCCCACTACCTGTGAGTTCAAGTCCATC

540

QY

181

TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp

200

DB

541

TACATGGCCAGAGGCCCGTGACGCTGCCGGCTACTACTACGTGACGCCAAGCTGAC

600

QY

201

IleThrSerHisAsnGluAspTyrThrIleValGlnTyrGluArgThrGluGlyArg

220

DB

601

ATCACCTCCCAACAAGAGACTACACCATCTGAGAGCAGTACGAGCGCACCGAGGGCCGC

660

QY

221

HisHisLeuPheLeu

225

DB

661

CACCACCTGTTCTTG

675

RESULT 55

ADW43589

ID

ADW43589

standard; DNA; 4200 BP.

AC

XX

ADW43589;

DT

XX

07-APR-2005

(first entry)

DE

XX

UBCP-lox-miRMSOD2-EGFP-lox-RFP.

KW

XX

ss; hairpin; Neuroprotective; shRNA; miRNA; aberrant gene function; gain-of-function mutation; neurological disease.

OS

XX

Synthetic.

PN

XX

WO2005007877-A2.

PD

XX

27-JAN-2005.

PF

XX

19-JUL-2004; 2004WO-US023789.

PR

XX

18-JUL-2003; 2003US-0488510P.

PA

XX

(UYMA-) UNIV MASSACHUSETTS.

PI

XX

Xu Z, Xia X;

DR

XX

WPI; 2005-112883/12.

PT

XX

New construct comprises a U6 promoter operably linked to a small hairpin RNA (shRNA) encoding nucleic acid sequence, useful for regulating shRNA expression in cells or animals.

PS

XX

Example 6; SEQ ID NO 9; 88bp; English.

CC

XX

This sequence represents the construct UBCP-lox-miRMSOD2-EGFP-lox-RFP. In this construct, the Ubiquitin C promoter (UBCP) normally directs synthesis of the hairpin and EGFP. However, upon exposure to cre, the hairpin and EGFP re excised thus preventing hairpin expression, while the RFP gene is under the control of the UBCP. This sequence is a complementary vector to UBCP-lox-RFP-lox-miRMSOD2-EGFP. The construct of the invention comprises a U6 promoter operably linked to a shRNA encoding nucleic acid sequence. The construct further comprises a first loxP site upstream of the promoter and a second loxP site downstream of the shRNA encoding sequence, the loxP sites are in the same orientation so that the promoter and encoding sequences are excisable upon exposure to Cre. The shRNA encoding sequence comprises a first stem-encoding portion, a loop-encoding portion, and a second stem-encoding portion, the loop-encoding portion comprising a first loxP site operably linked to a transcription

CC

XX

termination signal upstream of a spacer DNA and a second loxP site, the loxP sites are in the same orientation so that the first loxP site, CC termination signal and spacer DNA are excisable upon exposure to Cre. The construct further comprises spacer DNA downstream of the shRNA encoding sequence, a second loxP site downstream of the spacer DNA, and a first loxP site within the loop-encoding portion of the shRNA encoding sequence, the loxP sites are in the same orientation so that the spacer DNA and second stem-encoding sequence are excisable upon exposure to Cre. The U6 promoter comprises a distal sequence element (DSE), a proximal sequence element (PSE), and a TATA box, operably linked, where the construct comprises a first loxP site downstream of the shRNA encoding sequence, and a second loxP site between the DSE and the PSE, the loxP sites are in the same orientation so that the shRNA encoding sequences and a portion of the promoter comprising the PSE and the TATA box are excisable upon exposure to Cre. The construct also comprises a first loxP site and a second loxP site, the sites are interrupted by spacer DNA, CC between the DSE and the PSE, the loxP sites are in the same orientation so that a loxP site and the spacer DNA are excisable upon exposure to CC Cre. The construct is useful for the treatment of a disease, which is CC caused by aberrant gene function. The disease is a dominant, gain-of- CC function mutation, e.g. a neurological disease. The constructs are useful CC for regulating shRNA expression in cells or animals. The constructs, CC methods, kits, and compositions are useful for investigating gene CC functions, both physiologic and pathologic, in specific cell groups and CC in specific ages, in normal and diseased pathways.

XX

SQ

Sequence 4200 BP; 889 A; 1094 C; 1265 G; 950 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:

8.85e-136

Length:

4200

Score:

1186.00

Matches:

219

Percent Similarity:

98.22%

Conservative:

2

Best Local Similarity:

97.33%

Mismatches:

4

Query Match:

97.69%

Indels:

0

DB:

14

Gaps:

0

US-10-006-922A-12 (1-225) x ADW43589 (1-4200)

QY

1

MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly

20

DB

3267

ATGGCCTCCTCCGAGAACGTATCATCACCAGTTCATGCGCTTCAAGGTGGCATGAGGCG

3326

QY

21

ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly

40

DB

3327

ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGGCCGCCCTTACGAGGGC

3386

QY

41

HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle

60

DB

3387

CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTTGGGACATC

3446

QY

61

LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro

80

DB

3447

CTGTCCCCCAGTTCAGTAGCAGCTCCAGGTGTACGTGAAGCACCCGCCGACATCCCC

3506

QY

81

AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu

100

DB

3507

GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG

3566

QY

101

AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr

120

DB

3567

GACGGCGCGGTGGAGACCGTGACCCAGGACTTCTCCCTGACGAGCGGCTTCATCTAC

3626

QY

121

LysValLysPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr

140

DB

3627

AAAGTGAAGTTATCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAGACACC

3686

QY

141

MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu

160

DB

3687

ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGTGTGAAGGGCGAG

3746

QY

161

IleHisLysAlaLeuLysLeuLysLeuAspGlyLysHisTyrLeuValGluPheLysSerIle

180

DB

3747

ACCCACAAGGCCCTGAAGCTGAAGAGCGGGCCACTACTGTGTGAGTTCAAGTCCATC

3806



Qy	181	TyTMeTAlaLySbLsProValGlnLeuProGlyTyTyTyTyValAspSerLySLeuA	200
Db	3807	TACATGGCCAGAGAGCCCGTGCAGCTGCCCGCTACTACTACTGAGCGCAAGCTGGAC	3866
Qy	201	IleThrSerHisAsnGluAspTyThrIleValGlnGlnTyArgTyrGluGlyArg	220
Db	3867	ATCACCCTCCACACGAGAGACTACACCATCGTGGAGCAGTACGAGCGACCGAGGCCGC	3926
Qy	221	HisHisLeuPheLeu	225
Db	3927	CACCACCTGTCTCTG	3941
RESULT 56			
ID	ADM43588	standard; DNA; 4300 BP.	
XX	AC	ADM43588;	
XX	DT	07-APR-2005 (first entry)	
XX	DE	UbCP-lox-RFP-lox-miRMSOD2-EGFP.	
XX	KM	ss; hairpin; Neuroprotective; shRNA; miRNA; aberrant gene function;	
XX	KW	gain-of-function mutation; neurological disease.	
XX	OS	Synthetic.	
XX	PN	WO2005007877-A2.	
XX	PD	27-JAN-2005.	
XX	PF	19-JUL-2004; 2004WO-US023789.	
PR	18-JUL-2003; 2003US-0488510P.		
XX	PA	(UYMA-) UNIV MASSACHUSETTS.	
XX	PI	Xu Z, Xia X;	
DR	WPI; 2005-112863/12.		
PT	New construct comprises a U6 promoter operably linked to a small hairpin		
PT	RNA (shRNA) encoding nucleic acid sequence, useful for regulating shRNA		
PT	expression in cells or animals.		
XX			
XX	Example 6; SEQ ID NO 8; 88bp; English.		
XX			
CC	This sequence represents the construct UbCP-lox-RFP-lox-miRMSOD2-EGFP. In		
CC	this construct, the Ubiquitin C promoter (UbCP) normally transcribes RFP		
CC	mRNA, which terminates at a polyA site. However, because the RFP gene is		
CC	flanked by two loxP sites, upon exposure to cre, the RFP gene is excised		
CC	and the construct is converted to UbCP-lox-miRMSOD2-EGFP. The resulting		
CC	recombined sequence transcribes a message that contains an intron with an		
CC	shRNA (small hairpin) against the mouse SOD2 gene and an EGFP gene. The		
CC	construct of the invention comprises a U6 promoter operably linked to a		
CC	shRNA encoding nucleic acid sequence. The construct further comprises a		
CC	first loxP site upstream of the promoter and a second loxP site		
CC	downstream of the shRNA encoding sequence, the loxP sites are in the same		
CC	orientation so that the promoter and encoding sequence comprises a first stem-		
CC	upon exposure to Cre. The shRNA encoding sequence comprises a first stem-		
CC	encoding portion, a loop-encoding portion, and a second stem-encoding		
CC	portion, the loop-encoding portion comprising a first loxP site operably		
CC	linked to a transcription termination signal upstream of a spacer DNA and		
CC	a second loxP site, the loxP sites are in the same orientation so that		
CC	the first loxP site, termination signal and spacer DNA are excisable upon		
CC	exposure to Cre. The construct further comprises spacer DNA downstream of		
CC	the shRNA encoding sequence, a second loxP site downstream of the spacer		
CC	DNA, and a first loxP site within the loop-encoding portion of the shRNA		
CC	encoding sequence, the loxP sites are in the same orientation so that the		
CC	spacer DNA and second stem-encoding sequence are excisable upon exposure		
CC	to Cre. The U6 promoter comprises a distal sequence element (DSE), a		
CC	proximal sequence element (PSE), and a TATA box, operably linked, where		

CC	the construct comprises a first loxp site downstream of the shRNA
CC	encoding sequence, and a second loxp site between the DSE and the PSE,
CC	the loxp sites are in the same orientation so that the shRNA encoding
CC	sequences and a portion of the promoter comprising the PSE and the TATA
CC	box are excisable upon exposure to Cre. The construct also comprises a
CC	first loxp site and a second loxp site, the sites are interrupted by
CC	spacer DNA, between the DSE and the PSE, the loxp sites are in the same
CC	orientation so that a loxp site and the spacer DNA are excisable upon
CC	exposure to Cre. The construct is useful for the treatment of a disease,
CC	which is caused by aberrant gene function. The disease is a dominant,
CC	gain-of-function mutation, e.g. a neurological disease. The constructs
CC	are useful for regulating shRNA expression in cells or animals. The
CC	constructs, methods, kits, and compositions are useful for investigating
CC	gene functions, both physiologic and pathologic, in specific cell groups
CC	and in specific ages, in normal and diseased pathways.
XX	
SQ	Sequence 4300 BP; 906 A; 1128 C; 1297 G; 967 T; 0 U; 2 Other;
Alignment Scores:	
Pred. No.:	9,15e-136
Score:	1186.00
Percent Similarity:	98.22%
Best Local Similarity:	97.33%
Query Match:	97.69%
DB:	14
	Gaps: 0
US-10-006-922A-12 (1-225) x ADM43588 (1-4300)	
QY	1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db	1291 ATGGCCTCCTCCGAGAACGTATCACCAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 1350
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db	1351 ACCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGGCCCCCTACGAGGGC 1410
QY	41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60
Db	1411 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCCCTTGGAGACATC 1470
QY	61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
Db	1471 CTGTCCCCCAGTTCAGTACGAGCTCCAAAGTGTACGTGAAGCACCCCGCCGACATCCCC 1530
QY	81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
Db	1531 GACTACAGAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTATGAATTCGAG 1590
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db	1591 GACGGCGCGGTGGCGACCGTGACCAAGACTCTCCCTGCAGAGCGGCTGTCTCATCTAC 1650
QY	121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
Db	1651 AAGGTGAAGTTTCATGCGCGGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAGAACC 1710
QY	141 MetGlyTyrGluLysAsnThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
Db	1711 ATGGGCTGGAGGCTCCACCGAGCGCCTGTACCCCGCAGCGCGCTGTGAAGGGCAG 1770
QY	161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
Db	1771 ACCCACAAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC 1830
QY	181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
Db	1831 TACATGGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCGCCAAAGCTGAC 1890
QY	201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db	1891 ATCAACCTCCCAAGAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCGC 1950
QY	221 HisHisLeuPheLeu 225



Db 1951 CACCACCTGTTCTG 1965

RESULT 57

ADQ91220 ADQ91220 standard; DNA; 4488 BP.

AC ADQ91220;

DT 04-NOV-2004 (first entry)

DE Fusion construct of human DCX promoter and the DSRed2 gene SeqID 6.

KW human; neuronal marker; doublecortin; DCX; ds; DSRed2; chimeric; bcl-2; PTEN; dopaminergic determination factor; Nurr1; neuronal migration; neurophilin; SDF-1; NeuroD; BMP4; transgenic; neurogenic; neural stem cell activity; neuronal differentiation; neurological disorder; epilepsy; stroke; learning; memory skill; neurodegenerative disorder; Parkinson's; Huntington's; HIV-dementia; psychiatric disorder; tumorous disorder; brain malformation; vasotropic; anticonvulsant; cerebroprotective; antiparkinsonian; neuroprotective; KW nootropic; anti-HIV; regulatory sequence; proliferative neuronal determined cell; anti-apoptotic.

OS Homo sapiens.

OS Chimeric.

OS Unidentified.

PN WO2004067751-A1.

XX 12-AUG-2004.

PF 28-JAN-2004; 2004WO-EP000760.

PR 28-JAN-2003; 2003EP-00002027.

PA (UYRE-) UNIV KLINIKUM REGENSBURG.

PI Couillard-Despres S, Karl C, Kuhn H, Aigner L;

XX WPI; 2004-581007/56.

PT Use of a regulatory sequence for the early, transient expression of a heterologous nucleotide sequence in proliferative neuronal determined cells.

PS Claim 20; SEQ ID NO 6; 270bp; English.

CC This invention relates to the novel use of regulatory sequences for the early, transient expression of a heterologous nucleotide sequences in proliferative neuronal determined cells. Specifically, it refers to the preparation of a recombinant vector comprising a heterologous gene sequence such as an anti-apoptotic gene e.g. bcl-2 or PTEN, a determination/differentiation gene such as the dopaminergic determination factor Nurr1, a gene capable of inducing neuronal migration e.g. neurophilin or SDF-1 or a transcription factor gene such as NeuroD or BMP4. The present invention describes non-human transgenic animals that have been transfected with such a vector that can be used for screening compounds in vitro to identify those with neurogenic properties capable of regulating neural stem cell activity, neurogenesis and/ or neuronal differentiation, which can subsequently be used in the development of appropriate pharmaceutical compositions. Furthermore, these transgenic animals can be used for the in vivo tracking of newly generated neurons, as well as for transplanted or migrating neuronal determined cells. CC Accordingly, such compositions can be used in gene therapy for neurological disorders e.g. epilepsy or stroke and for the enhancement of learning and/or memory skills caused by neurodegenerative disorders such as Parkinson's, Huntington's or HIV-dementia, as well as psychiatric disorders, tumorous disorders of the nervous system or a developmental or brain malformation. Hence, these compositions exhibit vasotropic, anticonvulsant, cerebroprotective, antiparkinsonian, neuroprotective, KW nootropic and anti-HIV activities. This polynucleotide sequence is the fusion construct DNA of the reporter gene DSRed2 under the control of the

CC human neuronal marker doublecortin (DCX) promoter, given in an CC exemplification of the invention.

XX Sequence 4488 BP; 1184 A; 1026 C; 973 G; 1305 T; 0 U; 0 Other;

SQ Alignment Scores:

Pred. No.: 9.72e-136 Length: 4488  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x ADQ91220 (1-4488)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluY 20  
DB 3572 ATGGCCTCCTCCGAGAACGTCATCACCGAGTTCATCGCCTTCAAGGTGCGCATGAGGCG 3631  
QY 21 ThrValasnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluY 40  
DB 3632 ACCGTGAACGCCACGAGTTCGAGTCGAGGGCGAGGGCGAGGGCGCCCTACGAGGGC 3691  
QY 41 HisAsnThrValLysLysLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 3692 CACAACACCGTGAAGCTGAAGGTGACCAAGCGCGCCCTGCTTGCCTGGACATC 3751  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 3752 CTGTCCCCCAGTTCAGTACGGCTCCAGGTGATCGTGAAGCACCCCGACATCCCC 3811  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 3812 GACTACAAGAACTGTCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 3871  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 3872 GACGGCGCGGTGGCGACCGTGACCCAGACTCTCCCTGCAGGACGGCTGCTCATCTAC 3931  
QY 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
DB 3932 AAGGTGAAGTTCATCGCGGTGAACCTTCCCGACGCGCCCGTGATGCAGAAAGAACCC 3991  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 3992 ATGGGCTGGAGCGCTCCACCGAGCGCTGTACCCCCGACGCGCGTGTGAAGGGCGAG 4051  
QY 161 IleHisLysAlaLeuLysLysLysLysAspGlyCysTyrLeuValGluPheLysSerIle 180  
DB 4052 ACCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACTGTGTGAGTTCAAGTCCATC 4111  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 4112 TACATGGCCAAAGAGCCGTGCAGCTGCCGCTACTACTACGTGAGCGCAAGCTGAC 4171  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluLysArg 220  
DB 4172 ATCACTCCCAAGAGAGACTACACCATCGTGAAGACAGTACGAGCGCACGAGGGCGCG 4231  
QY 221 HisHisLeuPheLeu 225  
DB 4232 CACCACCTGTTCTG 4246  
RESULT 58  
ADL35208 ADL35208 standard; DNA; 6706 BP.  
XX AC ADL35208;  
XX AC 03-JUN-2004 (first entry)  
DT Plasmid pLL3.2 nucleotide sequence SEQ ID NO:4.  
DE  
XX



KW	lentiviral vector; functional packaging signal; multiple cloning site;
KW	antimicrobial; gene therapy; infection; plasmid; pL3.2; circular; gene;
KW	ds.
XX	
OS	Human immunodeficiency virus 1.
OS	Synthetic.
XX	
FH	Key
FT	promoter
FT	Location/Qualifiers
FT	212. .816
FT	/tag= a
FT	/note= "CMV promoter/enhancer 1"
FT	835. .1509
FT	/tag= b
FT	/note= "5', HIV R-US-del gag (HIV NL4-3/454-1126) "
FT	1539. .2396
FT	/tag= c
FT	/note= "HIV RRE (HIV NL4-3/7622-8459) "
FT	2422. .2599
FT	/tag= d
FT	/note= "HIV Flap"
FT	2710. .2745
FT	/tag= e
FT	/note= "loxP"
FT	2772. .3452
FT	/tag= f
FT	/product= "dsRed2"
FT	3506. .3539
FT	/tag= g
FT	/note= "loxP"
FT	3594. .4183
FT	/tag= h
FT	/note= "WRE element"
FT	4203. .4892
FT	/tag= i
FT	/note= "3', SIN LTR"
FT	4895. .5568
FT	/tag= j
FT	/note= "pUC"
FT	5713. .6573
FT	/tag= k
FT	/product= "AmpR"
XX	
PN	WO2004022722-A2.
XX	
PD	18-MAR-2004.
XX	
PF	05-SEP-2003; 2003WO-US028111.
XX	
PR	06-SEP-2002; 2002US-0408558P.
PR	27-SEP-2002; 2002US-0414195P.
PR	21-NOV-2002; 2002US-0428039P.
XX	
PA	(MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX	
PI	Bear JE, Dillon CP, Robinson DA, Van Parijs L,
XX	
DR	WPI; 2004-315566/29.
XX	
PT	New lentiviral vector comprising a nucleic acid whose sequence includes a
PT	functional packaging signal, a multiple cloning site (MCS), and at least
PT	one additional element, useful for treating infectious diseases.
XX	
PS	Claim 43; SEQ ID NO 4; 197pp; English.
XX	
CC	The present invention describes a lentiviral vector comprising a nucleic
CC	acid whose sequence includes a functional packaging signal, a multiple
CC	cloning site (MCS), and at least one additional element selected from a
CC	second MCS, a second MCS into which a heterologous nucleic acid is
CC	inserted, an HIV FLAP element, an expression-enhancing
CC	posttranscriptional regulatory element, a target site for a site-specific
CC	recombinase, and a self-inactivating (SIN) LTR. Also described: (1) a
CC	lentiviral vector comprising the RNA polymerase III promoter; (2) a
CC	collection of at least two lentiviral vectors; (3) a three plasmid

lentiviral expression system; (4) a four plasmid lentiviral expression system comprising the three plasmid lentiviral expression system, further comprising a four plasmid comprising a nucleic acid sequence that encodes Rev, operably linked to a promoter; (5) a cell comprising the lentiviral vector; (6) a cell comprising a provirus derived from the lentiviral vector; (7) a transgenic animal, at least some of whose cells contain a provirus derived from the lentiviral vector; (8) a method of creating a producer cell line; (9) a method of producing lentiviral particles; (10) a method of expressing a heterologous nucleic acid in a target cell; (11) a method for achieving controlled expression of a heterologous nucleic acid in a cell; (12) a method for expressing a transcript in a mammal in a cell type or tissue-specific manner; (13) a composition comprising the lentiviral vector, and a delivery agent that enhances delivery of the vector to cells; (14) a pharmaceutical composition comprising the lentiviral vector and a carrier; (15) methods of reversibly inhibiting or reducing expression of a target transcript in a cell, or in a mammal in a cell type or tissue-specific manner; (16) a method of treating or preventing infection by an infectious agent, or a disease or clinical condition; and (17) a kit comprising the lentiviral transfer plasmid and a packaging mix. The lentiviral vector has antimicrobial activity, and can be used in gene therapy. The lentiviral vector, compositions and methods are useful for treating infections. The present sequence represents the plasmid pLV3.2 nucleotide sequence, which is used in the exemplification of the present invention.

DB:	12	Gaps:	0
Query Match:	97.69%	Indels:	0
Best Local Similarity:	97.33%	Mismatches:	4
Percent Similarity:	98.22%	Conservative:	2
Score:	1186.00	Matches:	219
Pred. No.:	1.71e-135	Length:	6706
Alignment Scores:			
Sequence	6706 BP, 1755 A, 1638 C, 1722 G, 1591 T, 0 U, 0 Other,		

US-10-006-922A-12 (1-225) X ADL35208 (1-6706)

QY	1	MetArgSerSerLysAsnValIle	ysgluphemetArgpheLysValArgMetglu	1	20
		:::			
Db	2778	ATGGCTCTCCGAGAACGTCA	TACCGAGTTCA	TGCGCTTCAAGGTGCGCATGGAAGGC	2837
QY	21	ThrValasnGlyHisglupheglu	IlegluglygluGlyArgProTyrglu	1	40
Db	2838	ACCGTGAACGGCCACGAGT	TCGAGATCGAGGGCGAGGGCCGCCCTACGAGGGC		2897
QY	41	HisAsnThrValLysLeuLysVal	ThrLysglYglyProLeuProPheAlaTyrAspIle	1	60
Db	2898	CACAACACCGTGAAGCTGA	AGGTGACCAAGGGCGGCCCTTGCTTCGCTTGGGACATC		2957
QY	61	LeuSerProGlnPhegluTyr	glYserLysValTyrValLysHisProAlaAspIlePro	1	80
Db	2958	CTGTCCCCCAGTTCCAGTAC	GCGCTCCAAAGGTGTACGTGAAGCACCCCGGCATCCCC		3017
QY	81	AspTyrLysLeuSerPhePro	gluglyPheLysTyrGluArgValMetAsnPheglu	1	100
Db	3018	GACTACAGAAGCTGTCTT	CCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG		3077
QY	101	AspGlyglYValValThrVal	ThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	1	120
Db	3078	GACGGCGGCGTGGCGAC	CGGTGACCCAGACTCTCTCTGCAGGACGGCTGCTTCACTTAC		3137
QY	121	LysValLysPheIleglYVal	AsnPheProSerAspGlyProValMetGlnLysLysThr	1	140
Db	3138	AAGTGAAATTCA	TCCGGCGTGAACCTTCCCTCCGACGCGCCCGTGATGCAGAGAAGACC		3197
QY	141	MetGlyTyrGluAlaSerThr	GluArgLeuTyrProArgAspGlyValLeuLysglYglu	1	160
Db	3198	ATGGGCTGGAGGCTCCAC	CGAGCGCCTGTACCCCCCGACGGCGTGTGAAGGGCGAG		3257
QY	161	IleHisLysAlaLeuLysLeu	LysAspGlyglYHisTyrLeuValGluPheLysSerIle	1	180
Db	3258	ACCCACAAAGCCCTGAAG	CTGAAGACCGGCGGCACTA	CTGTGTGAGTTCAAGTTCATC	3317



Oy		161	TyrMecAlALysIlySProValGlnLeuProGIYTYrTYrTYrValAspSerLYsLeuAsp	200
Dd		3318	TACATGGCCAAAGAGCCCGTGACACTGCCCGGCTACTACTAAGTGAAGCCCAAGCTGCAC	3377
Oy		201	IleThrSerHisAsnGluAspTYrThrIleValGlnTYrGluArgThrglGlyArg	220
Dd		3378	ATCACCCTCCACAACGAGACTACACCATCGTGAGAGAGTAAGAGCGACGAGGGCCGC	3437
Oy		221	HISHisLeuPheLeu	225
Dd		3438	CACCACCTGTTCCTG	3452
		RESULT 59		
ID	ADR30815		standard; DNA; 7495 BP.	
XX	ADR30815;			
DT	04-NOV-2004	(first entry)		
DE	Zebrafish BDNF gene construct DNA #3.			
KW	Brain-derived neurotrophic factor; BDNF; promoter; zebrafish; ds.			
OS	Danio rerio.			
OS	Chimeric.			
XX	Unidentified.			
FH	Key	Location/Qualifiers		
FT	misc_feature	1..20		
FT		/tag= a		
FT		/note= "Vector DNA"		
FT	misc_feature	15..5091		
FT		/tag= b		
FT		/note= "Expression vector DNA"		
FT	misc_feature	21..1776		
FT		/tag= c		
FT		/note= "Zebrafish BDNF gene 5' flank region"		
FT	promoter	1777..1804		
FT		/tag= d		
FT		/note= "Zebrafish BDNF promoter gene"		
FT	exon	1805..2099		
FT		/tag= e		
FT		/note= "Exon 1c (5' UT) of Zebrafish BDNF gene "		
FT	misc_feature	2100..2122		
FT		/tag= f		
FT		/note= "Vector DNA"		
FT	misc_feature	2123..2800		
FT		/tag= g		
FT		/note= "Reporter vector DNA"		
FT	misc_feature	2801..2807		
FT		/tag= h		
FT		/note= "Linker DNA"		
FT	misc_feature	2808..5086		
FT		/tag= i		
FT		/note= "3' flank region"		
FT	misc_feature	5087..7495		
FT		/tag= j		
FT		/note= "Vector DNA"		
XX		US2004157294-A1.		
PN				
PD		12-AUG-2004.		
XX				
Pf		23-DEC-2003; 2003US-00742828.		
XX				
PR		30-NOV-2001; 2001US-0334079P.		
PR		29-NOV-2002; 2002US-00306737.		
XX				
PA		(USGO ) US DEPT VETERANS AFFAIRS.		
PI		Heinrich G, Huynh G;		
XX				

```

DR WPI; 2004-592772/57.
XX Nucleic acid construct useful for screening substances for their ability
PT to influence production of brain derived neurotrophic factor (BDNF),
PT comprises promoter of BDNF gene and red fluorescent reporter.
XX Claim 26; SEQ ID NO 5; 56bp; English.
XX The present invention provides an isolated brain-derived neurotrophic
CC factor (BDNF) gene promoter and its associated nucleic acid constructs.
CC The invention is useful for screening a number of substances for their
CC ability to influence the production of BDNF in living organisms. The
CC present sequence is zebrafish brain-derived neurotrophic factor (BDNF)
CC gene construct DNA. This sequence is used in the invention.
XX
SQ Sequence 7495 BP; 2249 A; 1599 C; 1659 G; 1988 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      2e-135          Length:      7495
Score:         1186.00        Matches:     219
Percent Similarity: 98.22%    Conservative: 2
Best Local Similarity: 97.33% Mismatches:   4
Query Match:    97.69%       Indels:      0
DB:             13           Gaps:        0

US-10-006-922A-12 (1-225) x ADR30815 (1-7495)

QY      1 MetArgSerSerLySAAnValIleLySGluPheMetArgPheLySValArgMetGluGly 20
Db      2123 ATGGCCTCCTCCGAAGACGTATCACCGAGTTCATGCGCTTCAAGGTGCCATGAGGGC 2182
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrgLugly 40
Db      2183 ACCGTGAACGGCCACGAGTTCCAGATCGAGGCCAGGGCCGAGGGCCGCCCTACGAGGGC 2242
QY      41 HisAbnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspile 60
Db      2243 CAACAACCCGTGAAGCTGAAGGTGACCAGAAGGGCGGCCCTTGCCCTTGCGTGGGACATC 2302
QY      61 LeuSerProGlnPheGlnTyrgLysSerLySValTyrgValLySHisProAlaAspilePro 80
Db      2303 CTGTCCCCCAGATTCCAATAAGGCTCCAAAGTGTACGTGAAGCACCCCAGCATCCCC 2362
QY      81 AspTyrgLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
Db      2363 GACTAACAAAGACTGTCTTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 2422
QY      101 AspgLygLyValValThrValThrgLnAspSerSerLeuGlnAspgLyCySpheileTyrg 120
Db      2423 GACGGCGCGCGTGGCGACCGGTGACCCAGACTCTTCCCTGCAGGACGGCGTGCTCATCTAC 2482
QY      121 LySValLySPheIleGlyValAsnPheProSerAspgLyProValMetGlnLySLeuThr 140
Db      2483 AAGGTGAAGTTTCATCGGCGTGAATCTCCCTCCGAGCGCCCGTGATGACAAAGAACCC 2542
QY      141 MetGlyTrpGluAlaSerThrgLuArgLeuTyrgProArgAspgLyValLeuLyegLyglu 160
Db      2543 ATGGGCTGGAGGCTCTCAACCGAGCGCTGTATCCCCCGCAGCGCGGTGCTGAAGGGCAG 2602
QY      161 IleHisLySAlaLeuLySLeuLySAspgLygLyHisTyrgLeuValGluPheLySSerile 180
Db      2603 ACCCAACAGGCCCTGAAGCTGAAGAGGACGGCGCCACTACTGTGTGAGTTCAAGTCTATC 2662
QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrgTyrgTyrgValAspSerLySLeuASP 200
Db      2663 TACATGGCCCAAGAGCCCGTGCAGCTGCCCGCTACTACTACGTGAGCGCAAAGCTGAGC 2722
QY      201 IleThrSerHisAsnGluAspTyrgThrIleValGluGlnTyrgLuArgThrgLuglyArg 220
Db      2723 ATCACTTCCCAACAAGAGACTACCATCTGTGAGAGCTACGAGCGCACCGAGGGCGCC 2782
QY      221 HisHisLeuPheLeu 225

```



Db 2783 CACCACTGTTCTTG 2797

RESULT 60

ADQ91222

ID ADQ91222 standard; DNA; 7616 BP.

XX

AC ADQ91222;

XX

DT 04-NOV-2004 (first entry)

XX

DE Expression vector phudCXpromodSRed2 DNA SeqID 8.

XX

KW human; neuronal marker; doublecortin; DCX; ds; DSRed2; expression vector; bcl-2; PTEN; dopaminergic determination factor; Nurr1; neuronal migration; neurophilin; SDF-1; NeuroD; BMP4; transgenic; neurogenic; neural stem cell activity; neuronal differentiation; neurodegenerative disorder; epilepsy; stroke; learning; memory skill; neurodegenerative disorder; Parkinson's; Huntington's; HIV-dementia; psychiatric disorder; tumorous disorder; brain malformation; vasotropic; anticonvulsant; cerebroprotective; antiparkinsonian; neuroprotective; neurotropic; anti-HIV; regulatory sequence;

KW proliferative neuronal determined cell; anti-apoptotic.

XX

OS Homo sapiens.

OS Synthetic.

OS Unidentified.

XX

PN WO2004067751-A1.

XX

PD 12-AUG-2004.

XX

PF 28-JAN-2004; 2004WO-EP000760.

XX

PR 28-JAN-2003; 2003EP-00002027.

XX

PA (UYRE-) UNIV KLINIKUM REGENSBURG.

XX

PI Coullard-Despres S, Karl C, Kuhn H, Aigner L;

XX

DR WPI; 2004-581007/56.

XX

PT Use of a regulatory sequence for the early, transient expression of a heterologous nucleotide sequence in proliferative neuronal determined cells.

PT

PS Claim 20; SEQ ID NO 8; 270pp; English.

XX

CC This invention relates to the novel use of regulatory sequences for the early, transient expression of a heterologous nucleotide sequences in proliferative neuronal determined cells. Specifically, it refers to the preparation of a recombinant vector comprising a heterologous gene sequence such as an anti-apoptotic gene e.g. bcl-2 or PTEN, a determination/differentiation gene such as the dopaminergic determination factor Nurr1, a gene capable of inducing neuronal migration e.g. neurophilin or SDF-1 or a transcription factor gene such as NeuroD or BMP4. The present invention describes non-human transgenic animals that have been transfected with such a vector that can be used for screening compounds in vitro to identify those with neurogenic properties capable of regulating neural stem cell activity, neurogenesis and/ or neuronal differentiation, which can subsequently be used in the development of appropriate pharmaceutical compositions. Furthermore, these transgenic animals can be used for the in vivo tracking of newly generated neurons, as well as for transplanted or migrating neuronal determined cells.

CC Accordingly, such compositions can be used in gene therapy for neurological disorders e.g. epilepsy or stroke and for the enhancement of learning and/or memory skills caused by neurodegenerative disorders such as Parkinson's, Huntington's or HIV-dementia, as well as psychiatric disorders, tumorous disorders of the nervous system or a developmental or brain malformation. Hence, these compositions exhibit vasotropic, anticonvulsant, cerebroprotective, antiparkinsonian, neuroprotective, neurotropic and anti-HIV activities. This polynucleotide sequence is the expression vector phudCXpromodSRed2, which comprises the reporter gene DSRed2 under the control of the human neuronal marker doublecortin (DCX)

CC promoter, given in an exemplification of the invention.

XX

SQ Sequence 7616 BP; 1916 A; 1854 C; 1822 G; 2024 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.: 2.05e-135 Length: 7616

Score: 1186.00 Matches: 219

Percent Similarity: 98.22% Conservative: 2

Best Local Similarity: 97.33% Mismatches: 4

Query Match: 97.69% Indels: 0

DB: 13 Gaps: 0

US-10-006-922A-12 (1-225) x ADQ91222 (1-7616)

QY 1 MetArgSerSerLybAsnValIleLySGluPheMetArgPheLybValArgMetGluGly 20

DB 3606 ATGGCCTCCTCCGAGACGTCATCACCGAGTTGCTTCAAGGTGCGCATGAGGGC 3665

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

DB 3666 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGAGGGCGCCCTACGAGGGC 3725

QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaAspIlePro 60

DB 3726 CACAACACCGTGAAGCTGAAGTGACCAAGGGCGGCCCTTGCCCTGGGACATC 3785

QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80

DB 3786 CTGTCCCCCAGTTCAGTACGGCTCCAGAGGTGATGAAAGCACCCGACATCCCC 3845

QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTyrGluArgValMetAsnPheGlu 100

DB 3846 GACTACAAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 3905

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB 3906 GACGCGCGCGTGGCGACCGTGACCCAGACTCTCTCTGCAAGACGGCTGCTCATCTAC 3965

QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140

DB 3966 AAGGTGAAGTTCATCGCGTGAATTCCCTCCGACGGCCCGTGATGCAAGAAAGACC 4025

QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyGlu 160

DB 4026 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGTAAGGGCGAG 4085

QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180

DB 4086 ACCACAAAGCCCTGAAGCTGAAGAGCGGCCCACTACTGTGAGTTCAGTCCATC 4145

QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200

DB 4146 TACATGGCCAAAGAGCCCGTGACGCTGCCGCTACTACTACGTGACGCCAAGCTGAC 4205

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 4206 ATCACTTCCCAACAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGC 4265

QY 221 HisHisLeuPheLeu 225

DB 4266 CACCACTGTTCTTG 4280

RESULT 61

ADL35211

ID ADL35211 standard; DNA; 7927 BP.

XX

AC ADL35211;

XX

DT 03-JUN-2004 (first entry)

XX

DE Plasmid pLU3.5 nucleotide sequence SEQ ID NO:7.

XX

KW lentiviral vector; functional packaging signal; multiple cloning site;



Accession	Feature	Location/Qualifiers
KW	antimicrobial; gene therapy; infection; plasmid; pL13.5; circular; gene;	
KW	ds.	
XX		
OS	Human immunodeficiency virus 1.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	promoter	212..816
FT		/tag= a
FT		/note= "CMV promoter/enhancer 1"
FT	misc_feature	835..1509
FT		/tag= b
FT		/note= "5' HIV R-US-del gag (HIV NL4-3/454-1126)"
FT	misc_feature	1539..2396
FT		/tag= c
FT		/note= "HIV RRE (HIV NL4-3/7622-8459)"
FT	misc_feature	2422..2599
FT		/tag= d
FT		/note= "HIV Flap"
FT	promoter	2632..3841
FT		/tag= e
FT		/note= "Ubc promoter"
FT	misc_feature	3931..3966
FT		/tag= f
FT		/note= "LoxP"
FT	CDS	3993..4673
FT		/tag= g
FT		/product= "dsRed2"
FT	misc_feature	4727..4760
FT		/tag= h
FT		/note= "LoxP"
FT	misc_feature	4815..5404
FT		/tag= i
FT		/note= "WRB element"
FT	misc_feature	5424..6113
FT		/tag= j
FT		/note= "3' SIN LTR"
FT	rep_origin	6116..6789
FT		/tag= k
FT		/note= "pUC"
FT	CDS	6934..7794
FT		/tag= l
FT		/product= "AmpR"
XX		
PN	WO2004022722-A2.	
XX		
PD	18-MAR-2004.	
XX		
XX	05-SEP-2003; 2003WO-US028111.	
PR	06-SEP-2002; 2002US-0408558P.	
PR	27-SEP-2002; 2002US-0414195P.	
PR	21-NOV-2002; 2002US-0428039P.	
XX		
PA	(MASI ) MASSACHUSETTS INST TECHNOLOGY.	
XX		
PI	Bear JE, Dillon CP, Robinson DA, Van Parijs L;	
XX		
DR	WPI; 2004-315566/29.	
XX		
PT	New lentiviral vector comprising a nucleic acid whose sequence includes a functional packaging signal, a multiple cloning site (MCS), and at least one additional element selected from a	
PT	one additional element, useful for treating infectious diseases.	
PS	Claim 43, SEQ ID NO 7; 197pp; English.	
XX		
CC	The present invention describes a lentiviral vector comprising a nucleic acid whose sequence includes a functional packaging signal, a multiple cloning site (MCS), and at least one additional element selected from a	
CC	second MCS, a second MCS into which a heterologous nucleic acid is	
CC	inserted, an HIV FLAP element, an expression-enhancing	
CC	posttranscriptional regulatory element, a target site for a site-specific	
CC	recombinase, and a self-inactivating (SIN) LTR. Also described: (1) a	

lenticral vector comprising the RNA polymerase III promoter; (2) a collection of at least two lenticral vectors; (3) a three plasmid lenticral expression system; (4) a four plasmid lenticral expression system comprising the three plasmid lenticral expression system, further comprising a four plasmid comprising a nucleic acid sequence that encodes Rev, operably linked to a promoter; (5) a cell comprising the lenticral vector; (6) a cell comprising a provirus derived from the lenticral vector; (7) a transgenic animal, at least some of whose cells contain a provirus derived from the lenticral vector; (8) a method of creating a producer cell line; (9) a method of producing lenticral particles; (10) a method of expressing a heterologous nucleic acid in a target cell; (11) a method for achieving controlled expression of a heterologous nucleic acid in a cell; (12) a method for expressing a transcript in a mammal in a cell type or tissue-specific manner; (13) a composition comprising the lenticral vector, and a delivery agent that enhances delivery of the vector to cells; (14) a pharmaceutical composition comprising the lenticral vector and a carrier; (15) methods of reversibly inhibiting or reducing expression of a target transcript in a cell, or in a mammal in a cell type or tissue-specific manner; (16) a method of treating or preventing infection by an infectious agent, or a disease or clinical condition; and (17) a kit comprising the lenticral transfer plasmid and a packaging mix. The lenticral vector has antimicrobial activity, and can be used in gene therapy. The lenticral vector, compositions and methods are useful for treating infections. The present sequence represents the plasmid pL13.5 nucleotide sequence, which is used in the exemplification of the present invention.

Alignment Scores:	
Pred. No.:	2.17e-135
Score:	1186.00
Percent Similarity:	98.22%
Best Local Similarity:	97.33%
Query Match:	97.69%
DB:	12
Length:	7927
Matches:	219
Conservative:	2
Mismatches:	4
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x ADL35211 (1-7927)

QY 1 MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGly 20  
DB 3999 ATGGCCCTCCTCCGAGAACGTTCATCACCGAGTTTCATGCGCTTCAAGTGCGCATGAGGCG 4058  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 4059 ACCGTGAACGCCACGAGTTCAGATCGAGGGCGAGGGCGGGCGCCCTACGAGGGC 4118  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 4119 CACAAACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTGGGACATC 4178  
QY 61 LeuSerProGluPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 4179 CTGTCCCCCAGTTCAGTACGGCTCCAAGGTGACGTGAAGCACCCCGACATCCCC 4238  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 4239 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 4298  
QY 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
DB 4299 GACGGCGCGCTGGCGACCGTGACCCAGAACTCCTCCCTGCAGACGCGCTGCTCATCTAC 4358  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 4359 AAGGTGAAGTTTCATGCGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAAGAAGACC 4418  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 4419 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCCGCGAGCGCGTGCTGAAGGGCGAG 4478  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180



Db 4479 ACCCAAGGCCCTGAAGCTGAAGACGGCGCCACTACCTGTGTGAGTTCAAGTCCATC 4538  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 4539 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGAGCGCAAGCTGGAC 4598  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 4599 ATCATTCCCAACAAGAGACTACCATCTGTGAGAGAGTACGAGCGCACCGAGGGCCGC 4658  
Qy 221 HisHisLeuPheLeu 225  
Db 4659 CACCACCTGTTCCTG 4673  
RESULT 62  
ADM82791  
ID ADM82791 standard; cDNA; 9731 BP.  
XX  
AC ADM82791;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE DNA repair pathway related retroviral vector cDNA with CMV promoter.  
XX  
KW inducer; inhibitor; DNA repair pathway; anti-HIV; cytosstatic; virucide;  
KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;  
KW cancer; human adult T-cell leukaemia; lymphoma;  
KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;  
KW gene therapy; human; cyclic; circular; CMV promoter; ss.  
XX  
OS unidentified.  
XX  
PN WO2003089573-A2.  
XX  
PD 30-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010302.  
XX  
PR 05-APR-2002; 2002US-0370376P.  
XX  
PA (FISH/) FISHEL R A.  
PA (YODE/) YODER K E.  
XX  
PI Fishel RA, Yoder KE;  
XX  
DR WPI; 2003-854096/79.  
XX  
PT Screening for compounds that modulate a DNA repair pathway and/or  
PT retroviral integration, useful for treating retroviral infection,  
PT comprises determining the amount of a retroviral cDNA circularization in  
PT the presence of the test compound.  
XX  
PS Claim 73; SEQ ID NO 5; 89pp; English.  
XX  
CC The invention relates to a novel method for screening for inducers or  
CC inhibitors of a DNA repair pathway by contacting at least one component  
CC of a DNA repair pathway with a non-circularized retroviral cDNA in the  
CC presence and absence of a test compound, and determining whether  
CC circularization of the cDNA is increased or decreased in the presence of  
CC the test compound. The DNA repair pathway components have the following  
CC activities: anti-HIV, cytosstatic, virucide, antidiabetic, and  
CC neuroprotective. The method is useful for identifying compounds that  
CC modulate a DNA repair pathway and/or retroviral activity. The compound is  
CC used in manufacturing a pharmaceutical composition for the treatment of a  
CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-  
CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes  
CC or multiple sclerosis) or for increasing the efficiency of gene delivery  
CC in a gene therapy. This polynucleotide represents a retroviral cDNA  
CC sequence of the invention.  
XX  
SQ Sequence 9731 BP; 2444 A; 2412 C; 2548 G; 2327 T; 0 U; 0 Other;  
XX  
Alignment Scores:

Pred. No.: 2.89e-135 Length: 9731  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0  
US-10-006-922A-12 (1-225) x ADM82791 (1-9731)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 5496 ATGGCCTCCTCCGAGAGCTCATCACCGAGTTCTATGCGCTTCAAGGTGCGCATGAGGCG 5555  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGluGlyArgProTyrGluGly 40  
Db 5556 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGCCCTTACGAGGCG 5615  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 5616 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGGACATC 5675  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 5676 CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGTACGTGAAGCACCCCGCAGATCCCC 5735  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 5736 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAATTCGAG 5795  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 5796 GACGGCGCGCTGGCGACCGGTGACCCAGACTCTCCCTGCGAGACGGCTGTCATCTAC 5855  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 5856 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAGAGAAGACC 5915  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 5916 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 5975  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 5976 ACCCAACAAGGCCCTGAAGCTGAAGAGCGCGGCACATACTGTGAGATTCAAGTCCATC 6035  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 6036 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGCTACTACTACGTGAGCGCAAGCTGGAC 6095  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
Db 6096 ATCATTCCCAACAAGAGACTACCATCTGTGAGAGTACGAGCGCACCGAGGGCCGC 6155  
Qy 221 HisHisLeuPheLeu 225  
Db 6156 CACCACCTGTTCCTG 6170  
RESULT 63  
ADM82792  
ID ADM82792 standard; cDNA; 9782 BP.  
XX  
AC ADM82792;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE DNA repair pathway related retroviral vector cDNA with MSH2 promoter.  
XX  
KW inducer; inhibitor; DNA repair pathway; anti-HIV; cytosstatic; virucide;  
KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;  
KW cancer; human adult T-cell leukaemia; lymphoma;  
KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;  
KW gene therapy; human; cyclic; circular; MSH2 promoter; ss.  
XX



OS Unidentified.  
XX  
PN WO2003089573-A2.  
XX  
PD 30-OCT-2003.  
XX  
PF 04-APR-2003; 2003WO-US010302.  
XX  
PR 05-APR-2002; 2002US-0370376P.  
XX  
PA (FISH/) FISHEL R A.  
PA (YODE/) YODER K E.  
XX  
PI Fishel RA, Yoder KE;  
XX  
DR WPI; 2003-854096/79.  
XX  
PT Screening for compounds that modulate a DNA repair pathway and/or  
PT retroviral integration, useful for treating retroviral infection,  
PT comprises determining the amount of a retroviral cDNA circularization in  
PT the presence of the test compound.  
XX  
PS Claim 73; SEQ ID NO 6; 89pp; English.  
XX  
CC The invention relates to a novel method for screening for inducers or  
CC inhibitors of a DNA repair pathway by contacting at least one component  
CC of a DNA repair pathway with a non-circularized retroviral cDNA in the  
CC presence and absence of a test compound, and determining whether  
CC circularization of the cDNA is increased or decreased in the presence of  
CC the test compound. The DNA repair pathway components have the following  
CC activities: anti-HIV, cytosstatic, virucide, antidiabetic, and  
CC neuroprotective. The method is useful for identifying compounds that  
CC modulate a DNA repair pathway and/or retroviral activity. The compound is  
CC used in manufacturing a pharmaceutical composition for the treatment of a  
CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-  
CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes  
CC or multiple sclerosis) or for increasing the efficiency of gene delivery  
CC in a gene therapy. This polynucleotide represents a retroviral cDNA  
CC sequence of the invention.  
XX  
SQ Sequence 9782 BP; 2628 A; 2305 C; 2351 G; 2498 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,91e-135 Length: 9782  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADM82792 (1-9782)

QY 1 MetArGSeSerLySAsnValIleLySGluPheMetArGPheLySValAArgMetGluGly 20  
DB 13 ATGGCCTCCTCCGAGAACGTATCATCACCGAGTTCAATGCGCTTCAAGGTGCGCATGGAGGCG 72  
QY 21 ThrValAsnGlyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 73 ACCGTGAACGCCACAGAGTTGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC 132  
QY 41 HISAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 133 CACAACACCGTGAAGCTGAAGTGACCAAGGCGGCCCCCTCGCCTGGGACATC 192  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHISProAlaAspIlePro 80  
DB 193 CTGTCCCCCAAGTTCCAGTACGGCTCCAAAGGTGTACTGAAGCAACCCCGCCGACATCCC 252  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnDheGlu 100  
DB 253 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 312  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120

DB 313 GACGCGCGCTGGCGACCGTGAACCAAGACTCCTCCTGCGAGGACGGCTGCTCATCTAC 372  
QY 121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 373 AAGTGAAGTTCACTCGCGTGAATTCCCTCCGACGCGCCCGTGATGCAGAAGAAGACC 432  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 433 ATGGGTGGAGGCGCTCCACCGAGCGCTGTACCCTCCGCGACGGCGTGCTGAAGGGCGAG 492  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHISLyLeuValGluPheLySLeuIle 180  
DB 493 ACCCAAGGCGCTGAAGCTGAAGACGCGGCGCACTACTGCTGAGATTCAAGTCCATC 552  
QY 181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 553 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCTACTACTACTGAGCGCAAGCTGGAC 612  
QY 201 IleThrSerHISAsnGlyAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 613 ATCAGCTTCCCAACAAGAGACTACCATCGTGAGACAGTACGACCGCACCGAGGGCGCG 672  
QY 221 HisHisLeuPheLeu 225  
DB 673 CACCACCTGTTCCTG 687

RESULT 64  
AAA48742  
ID AAA48742 standard; cDNA; 898 BP.  
XX  
AC AAA48742;  
XX  
DT 15-SEP-2003 (revised)  
DT 19-SEP-2000 (first entry)  
XX  
DE Discosoma sp. "red" novel fluorescent protein drFP583 cDNA.  
XX  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW fluorescent labeling; ss.  
XX  
OS Discosoma sp; "red".

FH Key Location/Qualifiers  
FT CDS 93..770  
FT /product= "drFP583"  
FT /transl\_except= (pos:456..457,aa:Val)  
FT /transl\_except= (pos:470..473,aa:Val)  
FT /note= "The codon at position 456..457 has an apparent 1  
FT nucleotide deletion which alters the reading frame. The  
FT codon at position 470..473 has an apparent 1 nucleotide  
FT insertion which corrects the altered reading frame"

XX WO200034326-A1.  
PN 15-JUN-2000.  
XX  
PD 10-DEC-1999; 99WO-US029473.  
XX  
PF 11-DEC-1998; 98US-00210330.  
XX  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
DR WPI; 2000-423381/36.  
XX  
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
PS Claim 6; Page 73-74; 86pp; English.



XX The present sequence is the cDNA sequence of drFP583 from *Discosoma* sp.  
CC "red", a non-bioluminescent species of the Class Anthozoa. drFP583 is a  
CC full-length cDNA encoding a novel fluorescent protein (nfp). Fluorescent  
CC proteins can be used in fluorescent labeling, a useful tool for marking a  
CC protein, cell or organism of interest. Unlike other markers used in  
CC protein labeling, such as beta-galactosidase and luciferase, fluorescent  
CC proteins do not require an exogenous cofactor or substrate. Methods  
CC involving fluorescent proteins are also less laborious and less difficult  
CC to control than the traditional methods of fluorescent labeling, where a  
CC protein of interest is purified and then covalently conjugated to a  
CC fluorophore derivative. Novel fluorescent proteins isolated from species  
CC of the Class Anthozoa can be used as markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations compared to prior art fluorescent  
CC proteins such as green fluorescent protein. (Updated on 15-SEP-2003 to  
CC standardise OS field)

XX Sequence 898 BP; 273 A; 172 C; 240 G; 213 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.78e-136 Length: 898  
Score: 1184.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.53% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x AAAA8742 (1-898)

QY 1 MetArgSerSerLySAAnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 93 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAGTTCGTATGGAAGCA 152  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyValArgProTyrGluGly 40  
DB 153 ACGGTCAATGGGCACGAGTTGAATAGAAGCGCAAGAGAGGGGAGGCCATACGAAGGC 212  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 213 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGACCTTGCCATTGGCTGGATATT 272  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 273 TTGTCAACCAATTTCAATATGAGCAAGATATATGCAAGCACCTGCGACATACCA 332  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
DB 333 GACTATAAAAAGCTGTCAATTCCTGAAGATTAAATGGAAAGGTCATGAACTTGAA 392  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 393 GACGGTGGCGTCTTACTGTAAACCCAGATTCCAGTTTGCAGGATGGCTGTTTCATCTAC 452  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 453 AAGTCAAGTTCAATTGGCGTTGAACCTTCTCCGATGACCTGTTATGCAAAAGAAGACA 512  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 513 ATGGGCTGGGAACCAAGCACTAGCGCTTGTATCCTCGTATGGCGGTGTGAAAAGAGAG 572  
QY 161 IleHisLeuAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuThr 180  
DB 573 ATTCATAAGGCTCTGAAGCTGAAAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 632  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 633 TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 692  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 693 ATAACAAGCCACAACGAGAGACTATACATCGTTGAGCAGTATGAAAGAACCAGGAGCGC 752  
QY 221 HisHisLeuPheLeu 225  
DB 753 CACCATCTGTTCCTT 767  
RESULT 65  
AAD03615  
ID AAD03615 standard; cDNA; 898 BP.  
XX  
AC AAD03615;  
DT 11-SEP-2003 (revised)  
DT 19-JUN-2001 (first entry)  
XX  
DE Discosoma sp. red chromo/fluorescent protein, drFP583 (NFP-6) cDNA.  
XX  
KW Anthozoa; Chromoprotein; fluorescent protein; drFP583; NFP-6; sunscreen;  
KW analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; invivo marker; selective filter; ss.  
XX  
OS Discosoma sp; red.  
XX  
FH Key Location/Qualifiers  
FT CDS 93..770  
FT /\*tag= a  
FT /product= "Chromo/fluorescent protein, drFP583 (NFP-6) "  
FT /transl\_except= (pos:540..600, aa:Leu-Gly)  
FT /note= "Insertion of 60 bases alters the reading frame"  
XX  
PN WO200127150-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 13-OCT-2000; 2000WO-US028477.  
XX  
PR 14-OCT-1999; 99US-00418529.  
PR 15-OCT-1999; 99US-00418917.  
PR 15-OCT-1999; 99US-00418922.  
PR 19-NOV-1999; 99US-00444338.  
PR 19-NOV-1999; 99US-00444341.  
PR 09-DEC-1999; 99US-00457556.  
PR 09-DEC-1999; 99US-00457898.  
PR 09-DEC-1999; 99US-00458144.  
PR 09-DEC-1999; 99US-00458477.  
PR 10-DEC-1999; 99WO-US029405.  
PR 14-JUN-2000; 2000US-0211607P.  
PR 14-JUN-2000; 2000US-0211609P.  
PR 14-JUN-2000; 2000US-0211626P.  
PR 14-JUN-2000; 2000US-0211627P.  
PR 14-JUN-2000; 2000US-0211687P.  
PR 14-JUN-2000; 2000US-0211766P.  
PR 14-JUN-2000; 2000US-0211860P.  
PR 14-JUN-2000; 2000US-0211888P.  
PR 14-JUN-2000; 2000US-0212070P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Tersikh A;  
XX  
XX WPI; 2001-266409/27.  
DR P-PSDB; AAB00375.  
XX  
XX An Anthozoa chromo- or fluorescent protein (p1) present in an environment  
PT other than its natural environment, useful as a label in analyte  
PT detection assays and as a selectable marker in recombinant DNA  
PT applications.  
XX  
PS Claim 13; Fig 6; 69pp; English.  
XX  
CC The present sequence is a *Discosoma* sp. red chromo/fluorescent protein,  
CC drFP583 (NFP-6) cDNA. NFP-6 is present in an environment other than its  
CC natural environment and has an absorbance maximum ranging from 250nm to



CC 750nm and more usually from 540nm to 580nm and emission maximum ranging from 275nm to 775nm and more usually from 565 to 605nm. The chromoproteins or fluorescent proteins are useful as labels in analyte detection assays, as selectable markers in recombinant DNA applications, as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator and as in vivo markers in animals. They are also useful in sunscreens and as selective filters. (Updated on 11-SEP-2003 to standardise OS field)

SQ Sequence 898 BP; 273 A; 173 C; 240 G; 212 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.78e-136 Length: 898
Score: 1184.00 Matches: 219
Percent Similarity: 98.22% Conservative: 2
Best Local Similarity: 97.33% Mismatches: 4
Query Match: 97.53% Indels: 0
DB: 4 Gaps: 0

US-10-006-922A-12 (1-225) x AAD03615 (1-898)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 93 ATGAGGCTCTCCAAAGAATGTTATCAAGAGATTCAAGAGTTTAAGGTTCCATGGAAGGA 152
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 153 ACGGTCATGGGCGACGAGTTGAATAGAGCGAAGAGAGGAGCGCATACGAAGGC 212
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheIleTrpAspIle 60
DB 213 CACAATACCGTAAGACTTAAGTAACCAAGGGGGACCTTGCCATTGCTGGGATATT 272
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProIleAspIlePro 80
DB 273 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGTCAAGCACCCGCGCATATACCA 332
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100
DB 333 GACTATAAAAAGCTGTCATTCTCGAAGGATTTAATGGAAAGGTCATGAACTTTGA 392
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 393 GACGGTGGCGTCGTTACTGTAACCCAGGATTCAGTTGCAGAGATGGCTGTTCATCTAC 452
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB 453 AAGTCAAAGTTCATTGGCGTTGAACCTTCCCTCCGATGGAACCTGTATGCAAAAGACA 512
QY 141 MetGlyTyrGluIleAserThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB 513 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAATGCGCTGTGAAGAGAG 572
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 573 ATTCAATAGGCTGGAAGCTGAAAAGACGGGTGTCATTACCTAGTTGAATTCAAAAGTATT 632
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB 633 TACATGGCAAGAAGCCCTGTCAGCTACCAAGGCTACTATGTTGACTCCAAACTGGAT 692
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 693 ATTAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGATATGAAGAACCAGAGGACGC 752
QY 221 HisHisLeuPheLeu 225
DB 753 CACCATCTGTTCTT 767

RESULT 66
ID ADC24128
XX ID ADC24128 standard; DNA; 678 BP.
XX AC ADC24128;

XX 18-DEC-2003 (first entry)
DT Discosoma red fluorescent protein variant fast T1 DNA.
XX DE
XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;
XX fluorescent protein variant; transcription induction detection;
KW fluorescence energy resonance transfer; FRET; protein kinase;
KW protein phosphatase; ion indicator; ds; mutant; fast T1.
XX OS
XX Synthetic.
OS Discosoma.
XX US2003059835-A1.
XX PN
XX 27-MAR-2003.
XX PD
XX 10-APR-2002; 2002US-00121258.
XX PF
XX 26-FEB-2001; 2001US-00794308.
XX PR
XX 24-MAY-2001; 2001US-00866538.
XX PA
XX (TSIE/) TSIE R Y.
XX (CAMP/) CAMPBELL R E.
XX PI
XX Tsien RY, Campbell RE;
XX WIPI; 2003-743764/70.
XX DR
XX P-PSDB; ADC24129.
XX DR
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein
PT variant having a reduced propensity to oligomerize, useful for detecting
PT transcriptional activity.
XX TS
XX Example 1; SEQ ID NO 3; 67pp; English.
PS
XX The invention describes a polynucleotide sequence (I) encoding a
CC Discosoma red fluorescent protein (DsRed) variant having a reduced
CC propensity to oligomerise, comprising amino acid substitutions at the AB
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225
CC amino acids, given in the specification, where the substitutions result
CC in reduced propensity of the DsRed variant to form tetramers. (I) is
CC useful for detecting transcriptional activity by providing a host cells
CC containing a vector which comprises (i) operatively linked to an
CC expression control sequence, and an unit to assay the variant fluorescent
CC protein fluorescence, and assaying fluorescence of the variant
CC fluorescence is indicative of transcriptional activity. A polynucleotide
CC encoding a fusion protein is useful for the analysis of in vivo
CC localisation or trafficking of a polypeptide of interest. A polypeptide
CC marker is useful as markers to identify the location and amount of a
CC target protein produced, where the target protein is fused to the marker,
CC as a complement to or alternative for the green fluorescent protein or
CC its spectral variant, for detecting induction of transcriptions, in
CC applications involving fluorescence energy resonance transfer (FRET),
CC which detects events as the function of the movement of fluorescent
CC donors and acceptors towards or away from each other, for making
CC fluorescent sensors for protein kinase and phosphatase activities or
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the
CC presence of a molecule in a sample, for identifying a specific
CC interaction of a first and second molecule, for determining whether a
CC sample contains an enzyme or for determining the pH of the sample. (I) is
CC useful for identifying a region or condition that regulates the activity
CC of an expression control sequence. This sequence encodes Discosoma red
CC fluorescent protein variant fast T1.
XX SQ
XX Sequence 678 BP; 141 A; 225 C; 205 G; 107 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.13e-133 Length: 678
Score: 1160.00 Matches: 216
Percent Similarity: 97.33% Conservative: 3
Best Local Similarity: 96.00% Mismatches: 6



Query Match:	95.55%	Indels:	0
DB:	10	Gaps:	0
US-10-006-922A-12 (1-225) x ADC24128 (1-678)			
QY	1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20	
DB	1 ATGGCCTCCTCCGAGACGTCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGGC	60	
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40	
DB	61 TCCGTGAACGGCCACGAGTTCGAGTCGAGGGCGGAGGGCCGCCCTACGAGGGC	120	
QY	41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTpaSpIle	60	
DB	121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTGAGACATC	180	
QY	61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80	
DB	181 CTGTCCCGCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCGCGACATCCC	240	
QY	81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100	
DB	241 GACTACAAGAAGCTGTCTTCCCGGAGGCTTCAAGTGGAGCGCGTGATGAATTCCAG	300	
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120	
DB	301 GACGGCGCGGTGTGACCGTGACCCAGGACTCCTCCGCGAGACGGCTCCTCATCTAC	360	
QY	121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr	140	
DB	361 AAGGTGAAGTTTCATCGGCGTGACTTCCCTCCGACGGCCCCGTAAATGCAAGAAGACT	420	
QY	141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160	
DB	421 ATGGGCTGGGAGGCTCTCCACCGAGCGCTGTACCCTCCGCGAGCGCGTGTGAAGGGCGAG	480	
QY	161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180	
DB	481 ATCCACAAGGCCCTGAAGCTGAAGGAGCGGCGGCACTACTGAGTGAAGTTCATC	540	
QY	181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200	
DB	541 TACATGGCCAAAGAGCCCGTGACGTGCGCCGCTACTACTACGTGAGACTCCAAGCTGAC	600	
QY	201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220	
DB	601 ATCACCTCCCAACAAGAGACTACCATCTGTGAGACGAGTACGAGCGCGCGAGGGCGCG	660	
QY	221 HisHisLeuPheLeu 225		
DB	661 CACCACTGTCTCTG 675		
RESULT 67			
ADL46207	standard, DNA; 678 BP.		
ID	ADL46207		
XX	ADL46207;		
XX	AC		
XX	DT 20-MAY-2004 (first entry)		
XX	DE Discosoma red fluorescent protein (DsRed) variant T1 coding sequence.		
XX	KW ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;		
KW	oligomerization; tetramerization; immunoassay; hybridization assay.		
XX	OS Discosoma sp.		
XX	XX		
FH	Key		
FT	CDS		
FT	location/Qualifiers		
FT	1..678		
FT	/*tag= a		
FT	/product= "DsRed variant T1 protein"		
.XX			

PN	W02003086446-A1.		
XX	23-OCT-2003.		
XX	09-APR-2003; 2003WO-US010879.		
PF	10-APR-2002; 2002US-00121258.		
XX	29-JUL-2002; 2002US-00209208.		
XX	(REGC ) UNIV CALIFORNIA.		
PA	Tsien RY, Campbell RE, Baird GS;		
XX			
PI	WPI; 2003-845265/78.		
XX	P-PSDB; ADL46206.		
DR			
PT	New monomeric and dimeric Anthozoan fluorescent protein variants with		
PT	reduced propensity to oligomerize, and encoding polynucleotides, useful		
PT	in molecular biology, e.g. in immunoassays or in tracking protein		
PT	movement in cells.		
XX			
PS	Disclosure; SEQ ID NO 5; 166bp; English.		
XX			
CC	The invention relates to a polynucleotide sequence encoding a Discosoma		
CC	red fluorescent protein (DsRed) variant having a reduced propensity to		
CC	oligomerize. The protein variant comprises one or more amino acid		
CC	substitutions at the AB and/or AC interface(s) of the wild-type DsRed		
CC	sequence, where the substitutions result in reduced propensity of the		
CC	DsRed variant to form tetramers and where the variant displays detectable		
CC	fluorescence of at least one red wavelength. The composition and methods		
CC	are useful in producing red fluorescent proteins having reduced		
CC	propensity for oligomerization, especially tetramerization. The protein		
CC	may be used in molecular biology and in other scientific applications,		
CC	such as in immunoassays or hybridization assays, or in tracking the		
CC	movement of proteins in cells. This sequence corresponds to the DsRed		
CC	coding sequence for the variant T1.		
XX			
SQ	Sequence 678 BP; 141 A; 225 C; 205 G; 107 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	1.13e-133	Length:	678
Score:	1160.00	Matches:	216
Percent Similarity:	97.33%	Conservative:	3
Best Local Similarity:	96.00%	Mismatches:	6
Query Match:	95.55%	Indels:	0
DB:	11	Gaps:	0
US-10-006-922A-12 (1-225) x ADL46207 (1-678)			
QY	1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20	
DB	1 ATGGCCTCCTCCGAGACGTCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGGC	60	
QY	21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40	
DB	61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC	120	
QY	41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTpaSpIle	60	
DB	121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTTCCCTTGAGACATC	180	
QY	61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80	
DB	181 CTGTCCCGCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCGCGACATCCC	240	
QY	81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100	
DB	241 GACTACAAGAAGCTGTCTTCCCGGAGGCTTCAAGTGGAGCGCGTGATGAATTCCAG	300	
QY	101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120	
DB	301 GACGGCGCGGTGTGACCGTGACCCAGGACTCCTCCTGCAAGACGGCTCCTCATCTAC	360	



QY	121	LYSVAlLYSPHeIleGlyValAsnPhEProSerASPGLyProValMetGlnLYbLYSThr	140
DB	361	AAGGTGAAGTTCAATCGCGCTGAACCTTCCCCCTCCGACGGCCCCCGTAATGCAGAGAAGACT	420
QY	141	MeG1YTrpG1uAlAsErtHrg1uArgLeuTYrProArgASPGLyValLeuLYSG1YGLu	160
DB	421	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCCGCGACGGCGTGTGAAGGGCGAG	480
QY	161	1LeH1bLYbAlAlEulYbLeuLYbASPGLYGLYH1STYrLeuValGluPhElybSerIle	180
DB	481	ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGGCACTACCTGTGTGAGTTCAAGTCCATC	540
QY	181	TyrMeAlAlYbLYbSProValGlnEulProG1YTYrTYrTYrValASPserLYbLeuASP	200
DB	541	TACATGGCCAGAAGACCCCGTGACGTGCCGGCTACTACTACGTGACTCCAAGCTGAC	600
QY	201	1LeThrSerH1bASng1uASPtyrThrIleValGluGlnTYrG1uArgThrGluGlyArg	220
DB	601	ATCACTCCCAACAAGAGGACTACCACTCTGTGAGCAGTACGAGCGCGCCAGGGCCGC	660
QY	221	H1bH1bLeuPhELeu	225
DB	661	CACCACTGTTCCTG	675

```

RESULT 68
ADN33980
ID      ADN33980 standard; DNA; 704 BP.
XX
AC      ADN33980;
XX
XX      01-JUL-2004 (first entry)
XX
DE      Mutant-type DsRED encoding sequence.

```

CW Cnidarian; fluorescence resonance energy transfer; FRET; wild-type DsRED;  
 KW mutant; mutein; ds.

OS D18C0808A BP.  
XX  
PN W02003054158-A2.

PD	03-JUL-2003.
XX	
PF	18-DEC-2002; 2002WO-US040539.

AA 19-DEC-2001; 2001US-0341723P.  
PR

PA (UYCH-) UNIV CHICAGO.

PI Bevis B, Glick B;  
XX  
DR WPI, 2003-569236/53.

PT Novel nucleic acid encoding a rapidly maturing chromo- or fluorescent  
PT mutant of a Chl<sub>a</sub>idian chromo- or fluorescent protein or its mutant,  
PT useful for applications involving chromo- or fluorescent proteins.

PS Claim 7; SEQ ID NO 3; 65bp; English.

AA The present invention relates to nucleic acid that encodes a rapidly  
CC maturing chromo- or fluorescent mutant of a Chlarian chromo- or  
CC fluorescent protein or its mutant. The protein is useful in applications  
CC involving nucleic acid encoding a chromo- or fluorescent protein and is  
CC useful for producing a chromo and/or fluorescent protein which involves  
CC growing the cell, whereby the protein is expressed, and isolating the  
CC protein substantially free of other proteins. The protein is useful in  
CC applications involving chromo- or fluorescent protein and is useful as  
CC PCR primers, hybridization probes, etc. The expression cassettes are  
CC useful for synthesizing related proteins. The chromoproteins are useful  
CC as coloring agents which are capable of imparting color or pigment to a  
CC particular composition of matter e.g. food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g., animals and plants. The chromoproteins  
CC may also find use as labels in analyte detection assays, e.g. assays for

biological analytes of interest and as selectable markers in recombinant DNA applications, e.g. the production of transgenic cells and organisms. The fluorescent proteins find use in a variety of different applications, e.g. in fluorescence resonance energy transfer (FRET) applications, as biosensors in prokaryotic and eukaryotic cells, in applications involving the automated screening of arrays of cells expressing fluorescent reporting groups by using microscopic imaging and electronic analysis, as second messenger detectors, and in fluorescence activated cell sorting applications and as *in vivo* marker in animals. The fluorescent proteins also find use in protease cleavage assays. The proteins can also be used in assays to determine the phospholipid composition in biological membranes and as a fluorescent timer. The present sequence represents the mutant-type DsRED encoding sequence.

Sequence 704 BP; 145 A; 237 C; 212 G; 110 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	1,19e-133
Score:	1160.00
Percent Similarity:	97.33%
Best Local Similarity:	96.00%
Query Match:	95.55%
DB:	11
Length:	704
Matches:	216
Conservative:	3
Mismatches:	6
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) X ADN33980 (1-704)

**QY**            1 MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgAsnGluGly 20  
     ||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
**Db**          20 ATGGCCCTCCTCCGAGGACGTATCAAGGAGTTTCATCGGCTTCAAGGTGC GCATGGAGGCC 79

QY 21 ThrValaenclYhiBgIupheglulIeGluglYgluglYargProTYrglglY 40  
:::|||||  
DB 80 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCCGAGGGCCGCCCTACGAGGGC 13

QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
|||  
140 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCCCTTGCCTCGGAGCATC 199  
Db

QY 61 LeuSerProGlnPheGlnTyrGlySerIleValTyrValLysHisProAlaAspIlePro 80  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 200 CTGTCCCTCCAGTTCCAGTACGGCTCCAGGTGTACTGTAAGCACCCCCGCATCCCC 259

Qy 81 AsPTyrlYsLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
|||  
Db 260 GACTCAGAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGGAGCGCGTGAATCACTTCGAG 319  
|||

Qy 101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 320 GACGGCGGCTGGTGAACCGTGAACCCAGACTCTCCCTGCAGACGGCTCTTCACTTAC 379

QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
 180 AAGGTGAAGTTATCGAGCTGAACCTCCCTCCGACGGCCCCGTAATGCAGAGAAGACT 439

141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuYsgIyGlu 160

161 ILeHlalyaAlaleuLybLeuLybAaPglYglhIstYrLeuValGluPhelySeriIle 180

**QY**

181 TyrMetAlaIyLysPProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||||  
ECO TTTCTCCCAATACCCCCCTCAGCTGCCGGCATCTAAGTGGAATTCCAAGCTGGAC 619  
|||||

201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Qy	221	HisHisLeuPheLeu	225
Db	680	CACCACCTGTCTCTG	694



RESULT	69
ID	ADM13536
AD	ADM13536 standard; DNA; 678 BP.
XX	
AC	ADM13536;
XX	
DT	15-JUL-2004 (first entry)
XX	
DE	DNA encoding superfolding coral DsRed protein #1.
XX	
KW	GFP; superfolder mutant; ds; green fluorescent protein;
KW	directed-evolution; folding interference domain; reporter protein;
KW	mutant; DsRed; coral.
XX	
OS	Discosoma sp.
OS	Synthetic.
XX	
PN	US2004078148-A1.
XX	
PD	22-APR-2004.
XX	
PF	24-APR-2003; 2003US-00423688.
XX	
PR	24-APR-2002; 2002US-00132067.
XX	
PA	(WALD/) WALDO G S.
XX	
PI	Waldo GS;
XX	
DR	WPI; 2004-340059/31.
DR	P-PSDB; ADM13537.
XX	
PT	New modified green fluorescent protein, useful as a reporter in
PT	expression studies.
XX	
PS	Disclosure; SEQ ID NO 40; 46bp; English.
XX	

[illegible]



XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescent protein variant; transcription induction detection;  
KW fluorescence energy resonance transfer; FRET; protein kinase;  
KW protein phosphatase; ion indicator; mutant; dimer2; ds.  
XX  
OS Synthetic.  
OS Discosoma.  
XX US2003059835-A1.  
PN 27-MAR-2003.  
PD  
XX 10-APR-2002; 2002US-00121258.  
PF  
XX 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
PR  
XX (TSIE/) TSIE R Y.  
PA (CAMP/) CAMPBELL R E.  
PA  
XX Tsien RY, Campbell RB;  
PI  
XX WPI; 2003-743764/70.  
DR P-PSDB; ADC24131.  
DR  
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
PT  
XX Example 1; SEQ ID NO 5; 67bp; English.  
PS  
XX The invention describes a polynucleotide sequence (I) encoding a  
XX Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcriptions, in  
CC applications involving fluorescence energy resonance transfer (FRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma red  
CC fluorescent protein variant dimer2.  
XX  
SQ Sequence 681 BP; 145 A; 231 C; 205 G; 100 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.71e-129 length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservative: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADC24130 (1-681)

QY 3 SerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGlyThrVal 22  
|||:::|||||  
DB 10 TCCTCCGAGGACGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTTG 69  
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42  
|||||  
DB 70 AACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGCACCCAG 129  
QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62  
|||  
DB 130 ACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTGCGGTGGACATCCTGTCC 189  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
|||||  
DB 190 CCCAGTTCAGTACGAGCTCCAGGCGCTACGTGAAGCACCCCGCATCCCGACTAC 249  
QY 83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102  
|||||  
DB 250 AAGAAGCTGCTCTCCCGAGGCTTCAAGTGGAGCGCGTGATGAATTTCGAGGACGGC 309  
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
|||  
DB 310 GCGGTGTGACCGTGACCCAGAGCTCTCTCGAGAGCGCACCGCTGATCTACAAGGTG 369  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
|||  
DB 370 AAGTTCGCGGACCACTTCCCCCGAGCGGCCCGTATGTGAGAGAAGACCATGGGC 429  
QY 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162  
|||||  
DB 430 TGGAGGCTTCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAGATCCAC 489  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
:::|||||  
DB 490 CAGGCCCTGAAGCTGAAGACGCGGCCACTACTGTTGAGTTCAAGACCATCTACATG 549  
QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
|||||  
DB 550 GCCAAGAAGCCGTGCAAGTGCCTGCTACTACTGAGACACCAAGTGAACATCACC 609  
QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222  
|||||  
DB 610 TCCCAACAAGGAGTACACCATCTGGAACAGTACGAGCGCTCCGAGGGCGCCACACAC 669  
QY 223 LeuPheLeu 225  
DB 670 CTGTTCTTG 678

RESULT 71  
ADL46209  
ID ADL46209 standard; DNA; 681 BP.  
XX  
AC ADL46209;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Discosoma red fluorescent protein (DsRed) variant dimer2 coding sequence.  
XX ds; gene; red fluorescent protein; DsRed; fluorescence; red wavelength;  
KW oligomerization; tetramerization; immunoassay; hybridization assay.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..681  
FT /\*tag= a  
FT /product= "DsRed variant dimer2 protein"  
XX WO2003086446-A1.  
XX 23-OCT-2003.  
XX



PF 09-APR-2003; 2003WO-US010879.  
XX 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Tsien RY, Campbell RE, Baird GS;  
PI  
XX WPI; 2003-845265/78.  
DR P-PSDB; ADL46208.  
XX  
PT New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
XX  
XX Disclosure; SEQ ID NO 7; 166pp; English.  
PS  
XX  
XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC coding sequence for the variant dimer2.  
XX  
SQ Sequence 681 BP; 145 A; 231 C; 205 G; 100 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.71e-129 Length: 681  
Score: 1121.00 Matches: 207  
Percent Similarity: 95.96% Conservative: 7  
Best Local Similarity: 92.83% Mismatches: 9  
Query Match: 92.34% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46209 (1-681)

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluIleThrVal 22  
Db 10 TCCTCCGAGACGTGCATCAAAAGATTTCATGCGCTTCAAGGTGCGCATGGAGGCTCCGTG 69  
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluIleHisAsn 42  
Db 70 AACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGCAACCCAG 129  
QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62  
Db 130 ACCGCCAAGCTGAAGGTGACCAAGGGGGCCCCCTGCCCTTGGGACATCCTGTCC 189  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
Db 190 CCCCAgTTCCAGTACGGCTCCAAAGGGCGTACGTGAAGCACCCCGCGACATCCCCGACTAC 249  
QY 83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102  
Db 250 AAGAAgCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACCTTCAGAGCAGGC 309  
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
Db 310 GGCgTGTGACCGTGACCCAGACTCTCTCCCTGCAGAGACGGCAGCTGATCTACAAGGTG 369  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
Db 370 AAGTTCGCGGGCACCAACTTCCCCCGCAGCGGCCCGGTATGTGAGAAGAACCATGGGC 429

QY 143 TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHis 162  
Db 430 TGGAGGCTCCACCGAGCGCTGTACCCCCCGACGGCGTGCTGAAGGGCGAGATCCAC 489  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
Db 490 CAGGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGGAGGTTCAAAGACCAATCTACATG 549  
QY 183 AlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThr 202  
Db 550 GCCAAGAAGCCCGTGAGCTGCCCGGCTACTACTACGTGACACCAAGCTGACATCAC 609  
QY 203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIleArgHisHis 222  
Db 610 TCCCAACAACGAGACTACACCATCTGTGAACAGTACGAGCGCTCCGAGGGCCGCCACCA 669  
QY 223 LeuPheLeu 225  
Db 670 CTGTTCTCTG 678

RESULT 72  
AAA48748  
ID AAA48748 standard; cDNA; 678 BP.  
XX AAA48748;  
AC 19-SEP-2000 (first entry)  
DT  
XX  
XX Humanised anthozoa novel fluorescent protein drFP583/dmFP592 hybrid cDNA.  
DE Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
XX Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KM Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
XX Anthozoa sp; "red".  
OS Discosoma sp; "red".  
OS Synthetic.

Key Location/Qualifiers  
FT CDS 1..678  
FT /product= "drFP583/dmFP592 hybrid protein"

W0200034326-A1.

15-JUN-2000.  
10-DEC-1999; 99WO-US029473.  
PF 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
XX Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
XX WPI; 2000-423381/36.  
DR P-PSDB; AAY99838.  
DR  
XX

Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX

Claim 6; Page 77; 86pp; English.

XX  
XX The present sequence is the nucleotide sequence of humanised  
CC drFP583/dmFP592. Amplified wild-type coding region fragments from drFP583  
CC and dmFP592 were digested with restriction endonucleases and combined  
CC together to produce a hybrid construct. drFP583 is a full-length cDNA  
CC encoding a novel fluorescent protein (nFP) from Discosoma sp. "red", a  
CC non-bioluminescent species of the Class Anthozoa. The hybrid sequence was  
CC changed to optimise the codons for expression of the protein in mammalian  
CC cells. Fluorescent proteins can be used in fluorescent labeling, a useful  
CC tool for marking a protein, cell or organism of interest. Unlike other  
CC markers used in protein labeling, such as beta-galactosidase and  
CC luciferase, fluorescent proteins do not require an exogenous cofactor or



CC substrate. Methods involving fluorescent proteins are also less laborious  
CC and less difficult to control than the traditional methods of fluorescent  
CC labeling, where a protein of interest is purified and then covalently  
CC conjugated to a fluorophore derivative. Novel fluorescent proteins  
CC isolated from species of the Class Anthozoa can be used as markers for  
CC gene expression and protein localization studies, and in fluorescence  
CC resonance energy transfer (FRET) reactions. They may have improved  
CC properties and better suitability for larger excitations compared to  
CC prior art fluorescent proteins such as green fluorescent protein

XX  
SQ Sequence 678 BP; 145 A; 217 C; 218 G; 98 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.36e-128	Length:	678
Score:	1119.00	Matches:	206
Percent Similarity:	96.89%	Conservative:	12
Best Local Similarity:	91.56%	Mismatches:	7
Query Match:	92.17%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922A-12 (1-225) x AAA48748 (1-678)

```
OY      1 MetArgSerSerLyAsnValIleLysGluPheMetArgPheLyValArgMetGluGly 20
        |||
DB      1 ATGAGCTGCAGCAAGAAGCGTGATCAAGAGTTTCATGCGTTCAAGTCCGATGAGAGGC 60

OY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGly 40
        |||
DB      61 ACCGTGAACGGCCACGAGTTCGAGATCAAGGGCGAGGGCGCGCCCTACGAGGGC 120

OY      41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTrpAspIle 60
        |||
DB      121 CACTGCAGCGTGAAGCTCATGTGATCAACGAGGGCGCCCTCCCTTCGACATC 180

OY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80
        |||
DB      181 CTCAGCCCCCAGTTCCAGTACGCGCAGCAAGGTGTACGTGAAGCACCCCGGCACATCCCC 240

OY      81 AspTyrLyLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
        |||
DB      241 GACTACAAGAAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGTGATGAACCTTCGAG 300

OY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        |||
DB      301 GACGGCGCGCTGTGACCGTGAGCCAGACAGACGACCTCAAGACGGCTGTTCATCTAC 360

OY      121 LyValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
        |||
DB      361 GAGGTGAAGTTTCATCGCGCTGAACCTTCCCGACGACGGCCCGTGATGCAGCGCGGACCC 420

OY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLyGlyGlu 160
        |||
DB      421 CGGGCTGGAGGCCACGACGAGCGGCTCTACCCCGGACGCGCGTGTCAAGGGCGAC 480

OY      161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180
        |||
DB      481 ATCCACATGCGCCCTCGGCTCGAGGGCGGGCCACTACCTCGTGAAGTTCAAGACATC 540

OY      181 TyrMetAlaLyLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLyLeuAsp 200
        |||
DB      541 TACATGGCCAAAGAGCCCGTGACGCTCCCGGCTACTACTAGTGACAGCACTCGAC 600

OY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
        |||
DB      601 ATCACCAAGCCACAACGAGACTACACCATCTGTGAGCAGTACGAGCGGACCGAGGGCCGG 660

OY      221 HisHisLeuPheLeu 225
        |||
DB      661 CACCACCTCTTCTTC 675

RESULT 73
AAD46280
ID      AAD46280 standard; DNA; 678 BP.
```

XX  
AC AAD46280;  
XX  
DT 27-DEC-2002 (first entry)  
XX  
DE Discosoma sp. drFP583/dmFP592 (6/9Q) protein encoding DNA.

KW Fluorescent protein; chromoprotein; protease cleavage assay; filter;  
KW fluorescence activated cell sorting application; fluorescent timer;  
KW biosensor; fluorescence resonance energy transfer application; FRET;  
KW colouring agent; recombinant DNA application; analyte detection assay;  
KW sunscreen; second messenger detector; drFP583 protein; dmFP592 protein;  
KW gene; ds.

OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /\*tag= a  
FT /product= "drFP583/dmFP592 (6/9Q) protein"

XX  
PN WO200268459-A2.

XX  
PD 06-SEP-2002.

XX  
PF 20-FEB-2002; 2002WO-US005749.

XX  
PR 21-FEB-2001; 2001US-0270983P.

XX  
PR 04-DEC-2001; 2001US-00006922.

XX  
PA (CLON-) CLONTECH LAB INC.

XX  
PI Lukyanov S, Lukyanov K, Yanushovich Y, Savitsky A, Fradkov A;

XX  
DR MPI; 2002-691654/74.

XX  
DR P-PSDB; AAE28835.

XX  
PT New nucleic acid encoding a non-aggregating chromo- or fluorescent mutant  
PT of an aggregating Chidarian chromo- or fluorescent protein or mutant for  
PT analyte detection assays or fluorescence activated cell sorting  
PT applications.

XX  
PS Disclosure; Page 72; 80pp; English.

XX  
CC The invention relates to nucleic acid molecules encoding non-aggregating  
CC chromo/fluorescent proteins and their mutants. Chromo/fluoro proteins are  
CC useful in analyte detection assays, as colouring agents, as markers in  
CC recombinant DNA applications, as sunscreens or filters, in fluorescence  
CC resonance energy transfer (FRET) applications, as biosensors in  
CC prokaryotic and eukaryotic cells, in screening assays, as second  
CC messenger detectors, in fluorescence activated cell sorting applications,  
CC in protease cleavage assays or as fluorescent timers. The present  
CC sequence is a DNA encoding Discosoma sp. drFP583/dmFP592 (6/9Q) protein  
CC of the invention

XX  
SQ Sequence 678 BP; 145 A; 217 C; 218 G; 98 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.36e-128	Length:	678
Score:	1119.00	Matches:	206
Percent Similarity:	96.89%	Conservative:	12
Best Local Similarity:	91.56%	Mismatches:	7
Query Match:	92.17%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922A-12 (1-225) x AAD46280 (1-678)

```
OY      1 MetArgSerSerLyAsnValIleLysGluPheMetArgPheLyValArgMetGluGly 20
        |||
DB      1 ATGAGCTGCAGCAAGAAGCGTGATCAAGAGTTTCATGCGTTCAAGTCCGATGAGAGGC 60

OY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGly 40
        |||
```



Db 61 ACCGTGAACGGCCACGAGTTGAGATCAAGGGCGAGGGCGAGGGCCGCTTACGAGGC 120  
QY 41 H1sAenThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACTGCAGCGGTGAAGCTCATGGTACCAAGGGGGGGCCCCCTCCCTTCGCTTCGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 CTCAGCCCCCAGTTCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 240  
QY 81 AspTyrIysIysLeuSerPheProGlnGlyPheIysTyrGlnArgValMetAsnPheGlu 100  
Db 241 GACTACAAAGAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGGTGATGAACCTTCAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGCTGTGACCGGTGAGCCAGACAGACGCTCAAGGACGGCTGCTTCATCTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
Db 361 GAGGTGAAGTTCATCGCGGTGAACCTTCCCGACGCGCCCGGTGATGCAGCGCGGACC 420  
QY 141 MetGlyTyrGlnAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuIysGlyIu 160  
Db 421 CGGGCGTGGAGGCCACGACGCGCGCTCTACCCCCGGGACGGCGTCTCAAGGGCGGAC 480  
QY 161 H1sIysAlaLeuIysLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180  
Db 481 ATCCACATGGCCCTCCGGCTCGAGGGCGGGCCCACTACCTCGTGAGTTCAAAGAGCATC 540  
QY 181 TyrMetAlaIysIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
Db 541 TACATGGCCAAAGAACCCCGTGAGCTTCCCGGCTACTACTAGTGAACAGCAAGCTTCAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGlnArgThrGluGlyArg 220  
Db 601 ATCACCAAGCCACAACGAGACTACCATCGTGGAGCAGTACGAGCGGACCGAGGGCCGG 660  
QY 221 HisIleuPheIu 225  
Db 661 CACCACTCTTCCTC 675

RESULT 74  
AAA48749  
ID AAA48749 standard; cDNA; 678 BP.  
XX  
AC AAA48749;  
XX  
DT 19-SEP-2000 (first entry)  
XX  
DE Humanised Anthozoa novel fluorescent protein drFP583/dmFP592-2G cDNA.  
XX  
KW Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;  
KM fluorescent labeling; hybrid; mutant; mutein; ss.  
XX  
OS Discosoma sp; "red".  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..678  
FT /product= "drFP583/dmFP592-2G mutant hybrid protein"  
XX  
PN WO200034326-A1.  
XX  
PD 15-JUN-2000.  
XX  
PF 10-DEC-1999; 99WO-US029473.  
XX  
PR 11-DEC-1998; 98US-00210330.  
PR 14-OCT-1999; 99US-00418529.  
XX  
PA (CLON-) CLONTECH LAB INC.  
.XX

PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Green G, Chen Y;  
PI Ding L;  
XX  
XX WPI; 2000-423381/36.  
DR P-PSDB; AAY99839.  
XX  
PT Novel fluorescent protein from non-bioluminescent Discosoma sp. red,  
PT useful for fluorescent labeling and as markers.  
XX  
XX Claim 6; Page 79; 86pp; English.  
PS  
XX  
XX The present sequence is derived from the nucleotide sequence of humanised  
CC hybrid drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel  
CC fluorescent protein (nfp) from Discosoma sp. "red", a non-bioluminescent  
CC species of the Class Anthozoa. The hybrid sequence was changed to  
CC optimise the codons for expression of the protein in mammalian cells. The  
CC present sequence encodes two substitutions, K15Q and T217S, which were  
CC incorporated by random mutagenesis. Fluorescent proteins can be used in  
CC fluorescent labeling, a useful tool for marking a protein, cell or  
CC organism of interest. Unlike other markers used in protein labeling, such  
CC as beta-galactosidase and luciferase, fluorescent proteins do not require  
CC an exogenous cofactor or substrate. Methods involving fluorescent  
CC proteins are also less laborious and less difficult to control than the  
CC traditional methods of fluorescent labeling, where a protein of interest  
CC is purified and then covalently conjugated to a fluorophore derivative.  
CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
CC can be used as markers for gene expression and protein localization  
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
CC They may have improved properties and better suitability for larger  
CC excitations compared to prior art fluorescent proteins such as green  
CC fluorescent protein  
XX  
SQ Sequence 678 BP; 143 A; 218 C; 218 G; 99 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.33e-127 length: 678  
Score: 1111.00 Matches: 204  
Percent Similarity: 96.89% Conservative: 14  
Best Local Similarity: 90.67% Mismatches: 7  
Query Match: 91.52% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x AAA48749 (1-678)  
QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
Db 1 ATGAGCTGCAGCAAGAACGTGATCAAGAGTTCATGCGGTTCCAGGTGCGGATGAGGGC 60  
QY 21 ThrValAsnGlyHisGlnPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTCAGATCAAGGGCGAGGGCGAGGGCCCTTACGAGGGC 120  
QY 41 HisAenThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACTGCAGCGGTGAAGCTCATGGTACCAAGGGGGGGCCCCCTCCCTTCGCTTCGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
Db 181 CTCAGCCCCCAGTTCAGTACGGCAGCAAGGTGTGCGTGAAGCACCCCGCCGACATCCCC 240  
QY 81 AspTyrIysIysLeuSerPheProGlnGlyPheIysTyrGlnArgValMetAsnPheGlu 100  
Db 241 GACTACAAAGAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGGTGATGAACCTTCAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGCTGTGACCGGTGAGCCAGACAGACGCTCAAGGACGGCTGCTTCATCTAC 360  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
Db 361 GAGGTGAAGTTCATCGCGGTGAACCTTCCCGACGCGCGCCCGGTGATGCAGCGCGGACC 420  
QY 141 MetGlyTyrGlnAlaSerThrGlnArgLeuTyrProArgAspGlyValLeuIysGlyIu 160



Db 421 CGGGGCTGGAGGCCAGCAGCGAGCGCTCTACCCCCGGAGCGCGTGTCAAGGGCGAC 480

QY 161 ILeHsIbSvAlAlLeuLySLeuLySaPrGlyGlyHisTyrLeuValGluPheLySLeu 180

Db 481 ATCCACATGGCCCTCCGCGCTCGAGGGCGGCGGCACTACTCGTGGAGTTCAGAGCATC 540

QY 181 TyrMetAlAlLySbSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200

Db 541 TACATGGCCAGAAGCCCGTGCAGCTCCCGGCTACTACTAGTGAGCAGCAGCTCGAC 600

QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

Db 601 ATCAACCAAGCCACACGAGGACTACCACTGTGAGCAGTACGAGCGGTCCAGGGCCGG 660

QY 221 HisHisLeuPheLeu 225

Db 661 CACCACTCTTCCTC 675

CC	XX	RESULT 75
CC	XX	AAAA8750
CC	XX	AAAA8750 standard; cDNA; 678 BP.
CC	XX	AAAA8750;
CC	XX	19-SEP-2000 (first entry)
CC	XX	Humanised Anthozoa novel fluorescent protein drFP583/dmFP592-Q3 cDNA.
CC	KW	Anthozoa; drFP583; fluorescent protein; non-bioluminescent organism;
CC	KW	fluorescent labeling; hybrid; mutant; mutein; ss.
CC	OS	Discosoma sp; "red".
CC	OS	Synthetic.
CC	FH	
CC	FY	Key Location/Qualifiers
CC	FT	CDS 1..678
CC	FT	/product= "drFP583/dmFP592-Q3 mutant hybrid protein"
CC	PN	WO200034326-A1.
CC	XX	
CC	PD	15-JUN-2000.
CC	XX	
CC	PF	10-DEC-1999; 99WO-US029473.
CC	PR	11-DEC-1998; 98US-00210330.
CC	PR	14-OCT-1999; 99US-00418529.
CC	XX	
CC	PA	(CLON-) CLONTech LAB INC.
CC	PI	Lukyanov SA, Pradkov AF, Labas YA, Matz MV, Green G, Chen Y;
CC	PI	Ding L;
CC	DR	WPI; 2000-423381/36.
CC	DR	P-PSDB; MAY99840.
CC	PT	
CC	PT	Novel fluorescent protein from non-bioluminescent Discosoma sp. red,
CC	XX	useful for fluorescent labeling and as markers.
CC	PS	Claim 6; Page 80-81; 86pp; English.
CC	XX	
CC	CC	The present sequence is derived from the nucleotide sequence of humanised
CC	CC	hybrid drFP583/dmFP592. drFP583 is a full-length cDNA encoding a novel
CC	CC	fluorescent protein (nFP) from Discosoma sp. "red", a non-bioluminescent
CC	CC	species of the Class Anthozoa. The hybrid sequence was changed to
CC	CC	optimise the codons for expression of the protein in mammalian cells. The
CC	CC	present sequence encodes three substitutions, K15Q, K83M and T217S, which
CC	CC	were incorporated by random mutagenesis. Fluorescent proteins can be used
CC	CC	in fluorescent labeling, a useful tool for marking a protein, cell or
CC	CC	organism of interest. Unlike other markers used in protein labeling, such
CC	CC	as beta-galactosidase and luciferase, fluorescent proteins do not require
CC	CC	an exogenous cofactor or substrate. Methods involving fluorescent
CC	CC	proteins are also less laborious and less difficult to control than the

CC traditional methods of fluorescent labeling, where a protein of interest  
CC is purified and then covalently conjugated to a fluorophore derivative.  
CC Novel fluorescent proteins isolated from species of the Class Anthozoa  
CC can be used as markers for gene expression and protein localization  
CC studies, and in fluorescence resonance energy transfer (FRET) reactions.  
CC They may have improved properties and better suitability for larger  
CC excitations compared to prior art fluorescent proteins such as green  
CC fluorescent protein  
XX  
SQ Sequence 678 BP; 142 A; 218 C; 218 G; 100 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	7.36e-127
Score:	1105.00
Percent Similarity:	96.44%
Best Local Similarity:	90.22%
Query Match:	91.02%
DB:	3
Length:	678
Matches:	203
Conservative:	14
Mismatches:	8
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) X AAA48750 (1-678)

Oy		1	MetArgSerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGly	20
Db		1	ATGAGCTGCAGCAAGAACGTGATCAAGAGTTTCATGCCGTTCCAGTGGATGGAGGCG	60
Oy		21	ThrValAsnGlyHisGluPheGlnIleGluGlyGluGlyArgProTyrgLugLy	40
Db		61	ACCGTGAACGGCCACGAGTTCGAGATCAAGGGCGAGGGCGGGCCTTAGAGGGC	120
Oy		41	HisAsnThrValLysLeuLysValThrlYsgLgLyProLeuProPheAlaTrpAspIle	60
Db		121	CACTGCAGCGTGAACTCATGGTGCACCAGGGCGGGCCCCCTCCCTTCGCTTCGACATC	180
Oy		61	LeuSerProGlnPheGlnTyrgLyserylYsvValTyrrValLysHisProAlaAspIlePro	80
Db		181	CTCAGCCCCCAGTTCAGTACGGCACCAAGGTGTACTGAAGCACCCCCCGACATCCCC	240
Oy		81	AspTyrrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db		241	GACTACATGAAGCTCAGCTTCCCAGAGGGCTTCAAGTGGAGCGGGTGAATTGAACCTTCGAG	300
Oy		101	AspGlyGlyValValThrValThrcGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db		301	GACGGCGGGCGTGGTGACCGGTGACGACGAGACAGCAGCCTCAAGAGACGGCTGCTTCATCTAC	360
Oy		121	LysValLysPheIleGlyValaAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db		361	GAGGTGAAGTTTCATCGGCGTGAAC TTC C C A G C G A C G G C C C G T G A T G A C G G C G G A C C	420
Oy		141	MetGlyTrpGluAlaSerThrcGlnArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db		421	CGGGGCTGGAGGCCAGCAGCAGCGGCTTA C C C C G G G A C G G C G T C T C A A G G G C G A C	480
Oy		161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHislYrLeuValGluPheLysSerIle	180
Db		481	ATCCACATGGCCCTCCGGCTCGAGGGCGGGCCACTACCTCGTGAGTTCAAGAGCATC	540
Oy		181	TyrrMetAlaLysLysProValGlnLeuProGlyTyrTyrrTyrrValaAspSerLysLeuAsp	200
Db		541	TACATGGCCAAGAAGCCCGTGCAGCTCCCGGCTACTACTACGTGAGACCAAGCTCGAC	600
Oy		201	IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrGluArgThrcGlnGlyArg	220
Db		601	ATCACCAGCCACAACGAGGACTACACCATCGTGAGACAGTACGAGCGGTCCGAGGGCGCG	660
Oy		221	HisHisLeuPheLeu	225
Db		661	CACCACTCTTCCTC	675
RESULT 76				
AAA28958				
ID	AAA28958	standard,	cdna;	876 BP.
XX				



AC AAA28958;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE Discosoma sp. "magenta" dmFP592 cDNA encoding a fluorescent protein.  
XX  
KW dmFP592; fluorescent protein; non-bioluminescent; label; marker;  
KM fluorescence resonance energy transfer reaction; FRET; ss.  
XX  
OS Discosoma sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 45..737  
FT /\*tag= a  
FT /product= "fluorescent\_protein"  
XX  
FT WO200034324-A1.  
XX  
PN 15-JUN-2000.  
XX  
PD 15-DEC-1999; 99WO-US029412.  
XX  
PF 10-DEC-1999; 99WO-US029412.  
XX  
PR 11-DEC-1998; 98US-00210330.  
PR 19-NOV-1999; 99US-00444341.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AP, Labas YA, Matz MV;  
XX  
XX WPI; 2000-423379/36.  
DR P-PSDB; AAY92894.  
DR  
PT Novel fluorescent protein from non-bioluminescent Discosoma species,  
PT useful for fluorescent labeling and as markers.  
XX  
XX Claim 7; Page 54-55; 60pp; English.  
PS  
XX  
CC This cDNA encodes a fluorescent protein (dmFP592) from Discosoma sp.  
CC "magenta", a non-bioluminescent species of the genera Anthozoa. The  
CC fluorescent colour is observed under usual white light or ultraviolet  
CC light. It has an absorbance maximum at 573 nm and an emission maximum at  
CC 593 nm. The polynucleotides of the invention can be used to produce the  
CC proteins recombinantly, and as a source of primers and probes for  
CC identifying related proteins. The fluorescent proteins have applications  
CC in fluorescent labeling studies, as fluorescent markers for gene expression and  
CC protein localization studies, and in fluorescence resonance energy  
CC transfer (FRET) reactions. They may have improved properties and better  
CC suitability for larger excitations, compared to prior art fluorescent  
CC proteins such as green fluorescent protein from Aequorea victoria, which  
CC has an excitation maximum at 395 nm, a second excitation peak at 475 nm  
CC and an emission maximum at 510 nm  
XX  
XX Sequence 876 BP; 256 A; 165 C; 232 G; 223 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 2.75e-124 Length: 876  
Score: 1085.50 Matches: 199  
Percent Similarity: 95.58% Conservative: 17  
Best Local Similarity: 88.05% Mismatches: 9  
Query Match: 89.42% Indels: 1  
DB: 3 Gaps: 1  
US-10-006-922A-12 (1-225) x AAA28958 (1-876)  
QY 1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 45 ATGAGTTGTTCCAAAGATGTGATCAAGAGTTTCATGAGGTTCAGGTTCTGATGGAAGG 104  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 105 ACGGTCAATGGGCAAGAGTTTGAATAAAGCGAAGGTGAAGGAGGCCCTTACGAAGGT 164  
QY 41 HisAenThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60

Db 165 CACTGTTCGTAAGCTTATGTGTAACCAAGGTGACCTTTGCCATTGTGCTTTGTATATT 224  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHisProAlaAspIlePro 80  
Db 225 TTGTACACCAATTTCAATGATGGAAGCAAGGTATATGTCAAAACACCTTGCCGACATACCA 284  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 285 GACTATAAAAAGCTTCATTCTCTGAGGATTTAAATGGAAAAGGTCATGAACTTTGAA 344  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
Db 345 GACGGTGGCGTTACTGTATCCCAAGATTCAATTGAAAGACGGCTGTTTCATCTAC 404  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 405 GAGGTCAAGTTCATTGGGGTGAACCTTCTTGATGACCTGTTATGACAGAGAGACA 464  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 465 CGGGCTGGGAAGCCAGCTCTGAGCGTTTGATCTCTGATGGGGTGTCAAAAGAGAC 524  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
Db 525 ATCCATATGCTCTGAGCTGGAAGAGGCGGCCATTACCTCGTTGAATTCAAAAGTATT 584  
QY 181 TyrMetAlaLySLeuPro--ValGlnLeuProGlyTyrTyrTyrValAspSerLySLeu 199  
Db 585 TACATGTAAAGAACCTTCAGTCAGTCAGTGCAGGCTACTATTATGTGACTCAAACTG 644  
QY 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 219  
Db 645 GATATGACGAGCCACAGAAATATACAGATCGTTGAGCAGATGAAAAAACCAAGGA 704  
QY 220 ArgHisHisLeuPheLeu 225  
Db 705 CGCCACCATCCGTTCAATT 722  
RESULT 77  
AAD03618  
ID AAD03618 standard; cDNA, 876 BP.  
XX  
XX AAD03618;  
AC  
XX  
DT 11-SEP-2003 (revised)  
DT 19-JUN-2001 (first entry)  
XX  
DE Discosoma sp. magenta chromo/fluorescent protein, dmFP592 (NFP-9) cDNA.  
KW Anthozoa; Chromoprotein; fluorescent protein; dmFP592; NFP-9; sunscreen;  
KM analyte detection assay; selectable marker; recombinant DNA application;  
KW biosensor; pH indicator; invivo marker; selective filter; ss.  
XX  
OS Discosoma sp; magenta.  
XX  
FH Key Location/Qualifiers  
FT CDS 45..737  
FT /\*tag= a  
FT /product= "Chromo/fluorescent protein, dmFP592 (NFP-9) "  
XX  
PN WO200127150-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 13-OCT-2000; 2000WO-US028477.  
XX  
PR 14-OCT-1999; 99US-00418529.  
PR 15-OCT-1999; 99US-00418917.  
PR 15-OCT-1999; 99US-00418922.  
PR 19-NOV-1999; 99US-00444338.  
PR 19-NOV-1999; 99US-00444341.  
PR 09-DEC-1999; 99US-00457556.



PR 09-DEC-1999; 99US-00457898.  
PR 09-DEC-1999; 99US-00458144.  
PR 09-DEC-1999; 99US-00458477.  
PR 10-DEC-1999; 99WO-US029405.  
PR 14-JUN-2000; 2000US-0211607P.  
PR 14-JUN-2000; 2000US-0211609P.  
PR 14-JUN-2000; 2000US-0211626P.  
PR 14-JUN-2000; 2000US-0211627P.  
PR 14-JUN-2000; 2000US-0211687P.  
PR 14-JUN-2000; 2000US-0211766P.  
PR 14-JUN-2000; 2000US-0211880P.  
PR 14-JUN-2000; 2000US-0211888P.  
PR 14-JUN-2000; 2000US-0212070P.  
XX (CLON-) CLONTECH LAB INC.  
XX PA  
XX PI Lukyanov SA, Fradkov AF, Labas YA, Matz MV, Tereskikh A;  
XX  
XX WPI; 2001-266409/27.  
XX DR P-PSDB; AAE00378.  
XX  
PT An Anthozoa chromo- or fluorescent proteicn (PI) present in an environment  
PT other than its natural environment, useful as a label in analyte  
PT detection assays and as a selectable marker in recombinant DNA  
PT applications.  
XX  
XX  
PS Claim 13, Fig 9; 69pp; English.  
XX  
CC The present sequence is a Discosoma sp. magenta chromo/fluorescent  
CC protein, dmp592 (NFP-9) cDNA. NFP-9 is present in an environment other  
CC than its natural environment and has an absorbance maximum ranging from  
CC 375nm to 775nm and more usually from 560nm to 590nm and emission maximum  
CC ranging from 395nm to 795nm and more usually from 580 to 610nm. The  
CC chromoproteins or fluorescent proteins are useful as labels in analyte  
CC detection assays, as selectable markers in recombinant DNA applications,  
CC as biosensors in prokaryotic and eukaryotic cells e.g. as pH indicator  
CC and as in vivo markers in animals. They are also useful in sunscreens and  
CC as selective filters. (Updated on 11-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 876 BP; 256 A; 165 C; 232 G; 223 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.75e-124 Length: 876  
Score: 1085.50 Matches: 199  
Percent Similarity: 95.58% Conservative: 17  
Best Local Similarity: 88.05% Mismatches: 9  
Query Match: 89.42% Indels: 1  
DB: 4 Gaps: 1

US-10-006-922A-12 (1-225) x AAD03618 (1-876)

QY 1 MetArGSeSerIyEaSnValIllelySGluPhemeTArGPhelysValArGMeTglUGly 20  
DB 45 ATGAGTTGTTCCAAGAAATGTGATCAAGAGTTCATGAGTTCAGGTTCTGTATGGAAGGA 104  
QY 21 ThrValAsnGlyHieGluPhheGluIlleGluGlyGluGlyGluGlyArGProTyrGluGly 40  
DB 105 ACGGTCAATGGGCAGCATTTGAATAAAGGCGAAGGTGAAGGAGGCTTACGAAGGT 164  
QY 41 HisAsnThrValIySleuLySValThrLySGlyGlyProLeuProPheAlaTPAspIle 60  
DB 165 CACTGTTCGTAAGCTTAAGCTTAAGCAAGGCTGACCTTGGCCATTGCTTTGATATT 224  
QY 61 leuSerProGlnPheGlnTyrGlySerIySValTyrValIySHisProAlaAspIlePro 80  
DB 225 TTGTCAACCAATTCAGTATGGAAGCAAGGTATATGCAACACCGCCGACATACCA 284  
QY 81 AspTyrIySlySleuSerPheProGluGlyPhelyeTTPGluArgValMetAsnPhheGlu 100  
DB 285 GACTATAAAAGCTGTCTATTCTGAGGAGATTAAATGGGAAAGGTCATGAACCTTGAA 344  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySphelIeTyr 120

DB 345 GACGGTGGCGTGTACTGTATCCCAAGATTCAGTTTGAAAGACGGCTGTTTCATCTAC 404  
QY 121 LysValIySphelIeGlyValAsnPhheProSerAspGlyProValMetGlnIySlyeThr 140  
DB 405 GAGGTCAAGTTCATTGGGGGTGAACCTTCTTGATGACCTGTATGCAGAGAGACA 464  
QY 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 465 CGGGGCTGGGAAGCCAGCTCTGAGCGTTTGATCCTCGTATGGGGTGTGAAGAGAGAC 524  
QY 161 IleHisLySAlaLeuLySleuLySAspGlyGlyHisTyrLeuValGluPhelySeriIle 180  
DB 525 ATCCATATGGCTCTGAGGCTGGAAGGCGGCATTACTTCGTTGAATTCAAAAGTATT 584  
QY 181 TyrMetAlaLySlySPro--ValGlnLeuProGlyTyrTyrTyrValAspSerIySleu 199  
DB 585 TACATGCTAAAGAGACCTTCAGTGCAGTTGCCAGGCTACTATATGTTGACTCCAAACTG 644  
QY 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 219  
DB 645 GATATGACGAGCCACAACGAAAGATTACAGACTGTTGAGCAGTATGAAAAACCCAGGGA 704  
QY 220 ArgHisHisLeuPhelu 225  
DB 705 CGCCACCATCCGTTCATY 722  
RESULT 78  
ACCA4649  
ID ACC44649 standard; DNA; 876 BP.  
XX  
AC ACC44649;  
XX  
DT 29-MAY-2003 (first entry)  
XX  
DE Red fluorescent protein (DsRed) encoding DNA SEQ ID NO:39.  
XX  
KW Chromosome-based platform; artificial chromosome; eukaryotic chromosome;  
KW att site; integrase; recombinase; Aces; gene therapy; transgenic animal;  
KW platform artificial chromosome expression system; gene; ds.  
XX  
OS Discosoma sp.  
OS Synthetic.  
XX  
PN WO200297059-A2.  
XX  
PD 05-DEC-2002.  
XX  
PF 30-MAY-2002; 2002WO-US017452.  
XX  
PR 30-MAY-2001; 2001US-0294758P.  
PR 21-MAR-2002; 2002US-0366891P.  
XX  
PA (CHRO-) CHROMOS MOLECULAR SYSTEMS INC.  
XX  
PI Perkins E, Perez C, Lindenbaum M, Greene A, Leung J, Fleming E;  
PI Stewart S, Shellard J;  
XX  
DR WPI; 2003-140461/13.  
DR P-PSDB; ABP96651.  
XX  
PT Novel eukaryotic chromosome comprising one or many att sites which  
PT permits site-directed integration in the presence of lambda-integrase,  
PT useful for site-specific recombination-directed integration of DNA of  
PT interest.  
XX  
PS Example 1; Page 191-192; 272pp; English.  
XX  
CC The present invention describes a eukaryotic chromosome (I) comprising  
CC one or several att sites, where an att site is heterologous to the  
CC chromosome, and permits site-directed integration in the presence of  
CC lambda-integrase. Also described: (1) a platform artificial chromosome  
CC expression system (Aces) (II) comprising several sites that participate  
CC in recombinase catalysed recombination; and (2) a method (M1) for



CC introducing a heterologous nucleic acid into a platform artificial  
CC chromosome. (I) can be used in gene therapy. (M1) is useful for  
CC introducing a heterologous nucleic acid molecule into a platform  
CC artificial chromosome, preferably an Aces. (II) is useful for producing a  
CC transgenic animal (e.g. a fish, insect, reptile, amphibian, arachnid, or  
CC mammal) by introducing (II) by cell fusion, lipid-mediated transfection  
CC by a carrier system, microinjection, microcell fusion, electroporation,  
CC microprojectile bombardment or direct DNA transfer into an embryonic  
CC cell, preferably a stem cell or an embryo. (II) comprises a heterologous  
CC nucleic acid that encodes a therapeutic product which is useful for  
CC making a library of Aces comprising random portions of a genome. ACC44612  
CC to ACC44732 and ABP96650 to ABP96657 represent sequences used in the  
CC exemplification of the present invention  
XX  
SQ Sequence 876 BP; 256 A; 165 C; 232 G; 223 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.75e-124 Length: 876  
Score: 1085.50 Matches: 199  
Percent Similarity: 95.58% Conservative: 17  
Best Local Similarity: 88.05% Mismatches: 9  
Query Match: 89.42% Indels: 1  
DB: 10 Gaps: 1

US-10-006-922A-12 (1-225) x ACC44649 (1-876)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 45 ATGAGTGTGTTCCAAGAATGTGATCAAGAGTTTCATGAGGTTCAAGGTTGATGGAAGGA 104  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 105 ACGGTCAATGGGCACGAGTTTGAATAAAGCGCAAGGTGAAGGAGCCCTTACGAAGGT 164  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 165 CACTGTTCGTAAGCTTATGGTAACCAAGGTTGACCTTGGCATTTGCTTTGATATT 224  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 225 TTGTCCACCAATTTCAGTATGAGCAAGGATATATGTCAACACCCCTGCCACATACCA 284  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrgLuarqValMetAsnPheGlu 100  
DB 285 GACTATATAAAGCTGTCAATTCTGAGGGAATTTAATGGAAAGGTCATGAACTTGAA 344  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 345 GACGGTCGTGGTACTGTATCCAAAGATTCCAGTTGAAAGCGGCTGTTTCATCTAC 404  
QY 121 LysValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
DB 405 GAGGTCAAGTTCAATGGGTGAACCTTCTTGATGACCTGTTATGACAGAGAGACA 464  
QY 141 MetGlyTrrgLuAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGly 160  
DB 465 CGGGGCTGGGAAGCCAGCTCTGAGCGCTTGTATCCTCGTATGGGGTGTGAAAGAGAC 524  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
DB 525 ATCCATATGCGCTCTGAGGCTGGAAGAGCGGCCCATTAACCTCGTTGAATTCAAAAGTATT 584  
QY 181 TyrMetAlaLySLeuPro--ValGlnLeuProGlyTyrTyrTyrValAspSerLySLeu 199  
DB 585 TACATGTGTAAGAAGCCTTCAGTGCAGTTGCCAGGCTACTATTATGTGACTCCAAACTG 644  
QY 200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGly 219  
DB 645 GATATGACGAGCCACAACGAAGATTACACAGTCGTGAGCAGTATGAAAAAACCCAGGGA 704  
QY 220 ArgHisHisLeuPheLeu 225  
DB 705 CGCCACCATCCGTTCAATT 722

RESULT 79  
ADCC24132  
ID ADC24132 standard; DNA; 678 BP.  
XX  
AC ADC24132;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
XX Discosoma red fluorescent protein variant mRFP1 DNA.  
DE  
XX Discosoma red fluorescent protein; DsRed; AB interface; AC interface;  
KW fluorescence energy resonance transfer; FRRET; protein kinase;  
KW protein phosphatase; ion indicator; mutant; mRFP1; ds.  
XX  
XX Synthetic.  
OS Discosoma.  
XX  
PN US2003059835-A1.  
XX  
PD 27-MAR-2003.  
XX  
PF 10-APR-2002; 2002US-00121258.  
XX  
XX 26-FEB-2001; 2001US-00794308.  
PR 24-MAY-2001; 2001US-00866538.  
XX  
PA (TSIE/) TSIE N R Y.  
PA (CAMP/) CAMPBELL R E.  
XX  
PI Tsien RY, Campbell RE;  
PI  
XX WPI; 2003-743764/70.  
DR P-PSDB; ADC24133.  
DR  
XX Novel polynucleotide sequence encoding Discosoma red fluorescent protein  
PT variant having a reduced propensity to oligomerize, useful for detecting  
PT transcriptional activity.  
PT  
XX  
XX Disclosure; SEQ ID NO 7; 67bp; English.  
PS  
XX  
XX The invention describes a polynucleotide sequence (I) encoding a  
CC Discosoma red fluorescent protein (DsRed) variant having a reduced  
CC propensity to oligomerize, comprising amino acid substitutions at the AB  
CC and/or AC interfaces of the wild-type DsRed sequence (S1) comprising 225  
CC amino acids, given in the specification, where the substitutions result  
CC in reduced propensity of the DsRed variant to form tetramers. (I) is  
CC useful for detecting transcriptional activity by providing a host cells  
CC containing a vector which comprises (I) operatively linked to an  
CC expression control sequence, and an unit to assay the variant fluorescent  
CC protein fluorescence, and assaying fluorescence of the variant  
CC fluorescent protein produced by (VII), where variant fluorescent protein  
CC fluorescence is indicative of transcriptional activity. A polynucleotide  
CC encoding a fusion protein is useful for the analysis of in vivo  
CC localisation or trafficking of a polypeptide of interest. A polypeptide  
CC marker is useful as markers to identify the location and amount of a  
CC target protein produced, where the target protein is fused to the marker,  
CC as a complement to or alternative for the green fluorescent protein or  
CC its spectral variant, for detecting induction of transcription, in  
CC applications involving fluorescence energy resonance transfer (FRRET),  
CC which detects events as the function of the movement of fluorescent  
CC donors and acceptors towards or away from each other, for making  
CC fluorescent sensors for protein kinase and phosphatase activities or  
CC indicators for ions and molecules such as Ca2+, Zn2+, for identifying the  
CC presence of a molecule in a sample, for identifying a specific  
CC interaction of a first and second molecule, for determining whether a  
CC sample contains an enzyme or for determining the pH of the sample. (I) is  
CC useful for identifying a region or condition that regulates the activity  
CC of an expression control sequence. This sequence encodes Discosoma red  
CC fluorescent protein variant mRFP1.  
XX  
SQ Sequence 678 BP; 149 A; 228 C; 208 G; 93 T; 0 U; 0 Other;



Alignment Scores:  
Pred. No.: 1,88e-116 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x ADC24132 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
DB	1	ATGGCCTCCTCCGAGACGTCATCAAGAGTTTCATGCGCTTCAAGTGGCATGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC	120
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
DB	121	ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCGCTGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
DB	181	CTGTCCCTCAGTTCAGTACGGCTCCAGAGGCTACGTGAAGCACCCCGACATCCCC	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	241	GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATCACTTCGAG	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	301	GACGGCGGGGTGACCGTGAACCCAGACTCTCCCTGCAGAGCGCGAGTTTCATCTAC	360
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr	140
DB	361	AAAGTGAAGCTGCGCGGACCAACTTCCCTCCGACGGCCCCGTATGCAAGAAGACC	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu	160
DB	421	ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCCGAGGACGGCGCCTGAAGGGCGAG	480
QY	161	IleHisValAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle	180
DB	481	ATCAAGATGAGGCTGAAGCTGAAGGACGGGCGCACTACGACGCCGAGGTCAAGACCAC	540
QY	181	TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	541	TACATGGCCAAGAGCCCGTGCAGCTGCCGCGCCTTACAAGACCGACATCAAGCTGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	601	ATCACCTCCCAACAAGAGACTACACCATCTGTGAACAAGTACGAGCGCGCGAGGGCGGC	660
QY	221	His 221	
DB	661	CAC 663	

RESULT 80  
ADL46211  
ID ADL46211 standard; DNA; 678 BP.

AC	ADL46211;
DT	20-MAY-2004 (first entry)
DE	Discosoma red fluorescent protein (DsRed) variant mRFP1 coding sequence.
XX	
XX	red fluorescent protein; DsRed; fluorescence; red wavelength;
KW	oligomerization; tetramerization; immunoassay; hybridization assay; gene;
KW	ds.
XX	

OS Discosoma sp.

XX	Key	Location/Qualifiers
FH	CDS	1..678
FT		/tag= a
FT		/product= "DsRed variant mRFP1 protein"

WO2003086446-A1.

23-OCT-2003.

09-APR-2003; 2003WO-US010879.

10-APR-2002; 2002US-00121258.

29-JUL-2002; 2002US-00209208.

(REGC ) UNIV CALIFORNIA.

Tsien RY, Campbell RE, Baird GS;

WPI; 2003-845265/78.

P-PSDB; ADL46210.

New monomeric and dimeric Anthozoan fluorescent protein variants with reduced propensity to oligomerize, and encoding polynucleotides, useful in molecular biology, e.g. in immunoassays or in tracking protein movement in cells.

Disclosure; SEQ ID NO 9; 166pp; English.

The invention relates to a polynucleotide sequence encoding a Discosoma red fluorescent protein (DsRed) variant having a reduced propensity to oligomerize. The protein variant comprises one or more amino acid substitutions at the AB and/or AC interface(s) of the wild-type DsRed sequence, where the substitutions result in reduced propensity of the DsRed variant to form tetramers and where the variant displays detectable fluorescence of at least one red wavelength. The composition and methods are useful in producing red fluorescent proteins having reduced propensity for oligomerization, especially tetramerization. The protein CC may be used in molecular biology and in other scientific applications, such as in immunoassays or hybridization assays, or in tracking the movement of proteins in cells. This sequence corresponds to the DsRed variant mRFP1 coding sequence.

SQ Sequence 678 BP; 149 A; 228 C; 208 G; 93 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1,88e-116 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46211 (1-678)

QY	1	MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly	20
DB	1	ATGGCCTCCTCCGAGGACGTCATCAAGAGTTTCATGCGCTTCAAGTGGCATGAGGGC	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTTCCCTGGAATC	120
QY	41	HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle	60
DB	121	ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCGCTGGACATC	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
DB	181	CTGTCCCTCAGTTCAGTACGGCTCCAAGGCTTACGTGAAGCACCCCGACATCCCC	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100



|||||  
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGCGCGCGTGTGACCGTGAACCACTCTCCCTGCAAGACGCGAGTTCATCTAC 360  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGCTGCGCGCACCACTTCCCTCCGACGCGCCCGTAATGCAAGAAAGACC 420  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGAGGCGCTCCACCGAGCGGATGTACCCCGAGACGCGCCCTGAAGGGCGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCGCACTACGACGCCGAGGTCAAGACCA 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCCAAAGAACCCGTGACGTGCGCGCGCTTACAAGACCGACATCAAGCTGAC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACAAGACGAGCACTACCATCTGTGAACAGTACGAGCGCGCGAGGGCGGC 660  
Qy 221 His 221  
Db 661 CAC 663

RESULT 81  
AEA54911  
ID AEA54911 standard; DNA; 2241 BP.

XX AEA54911,  
XX 25-AUG-2005 (first entry)  
XX Fluorescence-related EGFP-caspase-3 DEVD linker-mRFP(x2) fusion DNA.  
XX fluorescence; protein localization; cellular transport; organelle;  
XX mitochondria; ds; gene; enhanced green fluorescent protein; EGFP;  
XX caspase-3.  
XX unidentified.  
XX Chimeric.  
XX WO2005054464-A1.  
XX 16-JUN-2005.  
XX 03-DEC-2004; 2004WO-JP018437.  
XX 03-DEC-2003; 2003JP-00404472.  
XX 27-JAN-2004; 2004JP-00018344.  
XX (RIKE ) RIKEN KK.  
XX (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
XX Miyawaki A, Kogure T, Hama H, Kinjo M, Saito K, Karasawa S;  
XX Araki T;  
XX WPI; 2005-444979/45.  
XX P-PSDB; AEA54910.  
XX Novel fluorescent protein existing as monomer and isolated from *Fungia*  
XX and *Montipora* species, useful for labeling and analyzing location and  
XX movement of intracellular organelles.  
XX Example 10; SEQ ID NO 50; 218pp; Japanese.  
XX The invention relates to a novel fluorescent protein existing as a

CC monomer and comprising any one of SEQ ID No. 1-29 (odd SEQ ID numbers)  
CC given in the specification, or any one of the sequences chosen from SEQ  
CC ID No. 1-29 (odd SEQ ID numbers), with one or more amino acid  
CC substitutions, deletions and/or additions, and having fluorescent  
CC characteristics. The fusion protein of the invention may be useful for  
CC analyzing the location or movement of a protein in a cell in vivo and for  
CC labeling and analyzing the location or movement of intracellular  
CC organelles, such as mitochondria. The current sequence is that of the  
CC fluorescence-related enhanced green fluorescent protein (EGFP)-caspase-3  
CC DEVD linker-mRFP(x2) fusion DNA of the invention.

XX Sequence 2241 BP; 503 A; 748 C; 670 G; 320 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.01e-115 Length: 2241  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 14 Gaps: 0

US-10-006-922A-12 (1-225) x AEA54911 (1-2241)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGlyGly 20  
Db 805 ATGCGCTCTCCGAGGACGTATCAAGAGATTCAATGCGCTTCAAGGTGCGCATGAGGGC 864  
Qy 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGlyGly 40  
Db 865 TCCGTGAACGCGCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCGCCCTACAGGGGC 924  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 925 ACCCAAGACCGCCAAAGTGAAGTGACCAAGGGCGGCCCTTGCCCTGCGGACATC 984  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 985 CTGTCCCTCAAGTTCAGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGCAGATCCC 1044  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 1045 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 1104  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 1105 GACGCGCGCGTGTGACCGTGAACCAAGACTCTCTCTGCAAGACGCGGAGTTCATCTAC 1164  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 1165 AAGGTGAAGCTGCGCGCACCACTTCCCTCCGACGCGCCGTAATGCAAGAAAGACC 1224  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 1225 ATGGGCTGGAGGCGCTCCACCGAGCGGATGTACCCGAGAGACGGCCCTGAAGGGCGAG 1284  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1285 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCGCACTACGACGCCGAGGTCAAGACCA 1344  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1345 TACATGGCCAAAGAACCCGTGACAGCTGCGCGCGCTTACAAGACCGACATCAAGCTGAC 1404  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1405 ATCACTCCCAACAAGACGAGCACTACCATCTGTGAACAGTACGAGCGCGCGAGGGCGGC 1464  
Qy 221 His 221  
Db 1465 CAC 1467

RESULT 82  
ADQ59565



ID ADQ59565 standard; DNA; 675 BP.  
XX AC ADQ59565;  
XX DT 07-OCT-2004 (first entry)  
XX DE Discosoma monomerised red fluorescent protein mRFP1 DNA.  
XX KM optically detectable signal; protein fragment complementation assay; PCA;  
KM molecular interaction detection; biomolecular interaction detection;  
KM fluorescent protein; multi-colour PCA; drug discovery; target validation;  
KM high-throughput screening; high-content screening; pathway mapping;  
KM drug mechanism-of-action study; biosensor; diagnostic;  
KM red fluorescent protein; spectral shift; mRFP1; gene; ds.  
XX OS Discosoma.  
XX FH Key location/Qualifiers  
FT CDS 1..675  
FT /\*tag= a  
FT /product= "Discosoma monomerised red fluorescent protein"  
XX PN US2004137528-A1.  
XX PD 15-JUL-2004.  
XX PF 01-DEC-2003; 2003US-00724178.  
XX XX 02-FEB-1998; 98US-00017412.  
PR 07-FEB-2000; 2000US-00499464.  
PR 24-MAY-2002; 2002US-00154758.  
PR 29-JAN-2003; 2003US-00353090.  
PR 09-APR-2003; 2003US-0461133P.  
XX XX (MICH/) WATSON MICHNICK S W.  
PA (MACD/) MACDONALD M L.  
PA (LAME/) LAMERDIN J.  
PI Watson Michnick SW, Macdonald ML, Lamerdin J;  
XX DR WPI; 2004-533363/51.  
DR P-PSDB; ADQ59566.  
XX PT Composition useful in protein fragment complementation assays for drug  
PT discovery and high-throughput screening, comprising complementary  
PT fragments of protein or mutant protein, generating optically detectable  
PT signal when associated.  
XX PS Disclosure; SEQ ID NO 15; 34pp; English.  
XX XX The invention describes a composition (I) comprising complementary  
CC fragments of a protein or mutant protein, where the fragments generate an  
CC optically detectable signal when associated, and each of the mutant  
CC protein fragments is fused to a separate molecule. Also described are:  
CC protein fragment complementation assays (PCAs) for detection of molecular  
CC interactions, involving reassembling separate fragments from an optically  
CC detectable protein, and detecting the reassembly by units of  
CC reconstitution of activity of the optically detectable protein, where the  
CC reassembly of the fragments is operated by the interaction of molecular  
CC domains fused to each fragment, and is independent of other molecular  
CC processes; detecting biomolecular interaction, involving selecting an  
CC appropriate optically detectable protein, effecting fragmentation of the  
CC optically detectable protein such that the fragmentation results in  
CC reversible loss of protein function, fusing or attaching fragments of the  
CC optically detectable protein separately to other molecules, reassociating  
CC the protein fragments through interactions of the molecules that are  
CC fused or attached to the fragments, and detecting the resulting optical  
CC signal; designing and engineering of PCAs based on fluorescent protein;  
CC and a method and composition for the construction of multi-color PCAs.  
CC (I) is useful in PCAs and other assays for drug discovery, target  
CC validation, high-throughput screening, high-content screening, pathway  
CC mapping, drug mechanism-of-action studies, biosensors and diagnostics.  
CC (I) is useful for engineering different colour PCAs for a variety of

CC applications in biology and biotechnology. This sequence encodes a  
CC Discosoma monomerised red fluorescent protein from which PCA fragments  
CC can be isolated.  
XX SQ Sequence 675 BP; 147 A; 229 C; 207 G; 92 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 1.03e-115 Length: 675  
Score: 1015.00 Matches: 191  
Percent Similarity: 88.69% Conservative: 5  
Best Local Similarity: 86.43% Mismatches: 25  
Query Match: 83.61% Indels: 0  
DB: 12 Gaps: 0  
  
US-10-006-922A-12 (1-225) x ADQ59565 (1-675)  
  
QY 1 MetArgSerSerLyAsnValIleLySGluPhMetArgPheLyValArgMetGluGly 20  
DB 1 ATGGCCTCCTCCGAGACGTCATCAAGAGTTGCGCTTCAAGGTGCGCATGAGGGC 60  
  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGGGCCCTACGAGGGC 120  
  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGCCCTGCTTCCCTTGGCATC 180  
  
QY 61 LeuSerProGluPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAGTTCAGTACGGCTCCAAGCCTACGTGAAGCACCCGCGCATCCCC 240  
  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACTGGAAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTGC 300  
  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGCGCGGTGTGTGACCGTGACCGACGACTCCTCCCTGACGAGCGGCGAGTTCACTAC 360  
  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140  
DB 361 AAGGTGAAGCTGCGCGGCAACCACTTCCCTCCGACGGCCCGTAATGCAAGAAGACC 420  
  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCCACCGAGCGGATGTACCCTGAGGACGGCGCTGAAGGGCGAG 480  
  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180  
DB 481 ATCAAGATGAGCTGAAGCTGAAGGACGGCGGCACTACGACGCCGAGTCAAGACCACC 540  
  
QY 181 TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGACGTGCCCCGGCGCTACAAGACGACATCAAGCTGAC 600  
  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACCTCCCAACAGAGACTACCATCTGTGAACAAGTACGAGCGCGCGAGGGCCGC 660  
  
QY 221 His 221  
DB 661 CAC 663  
  
RESULT 83  
ADL46282  
ID ADL46282 standard; DNA; 678 BP.  
XX AC ADL46282;  
XX DT 20-MAY-2004 (first entry)  
XX XX Discosoma red fluorescent protein (DsRed) variant mRFP1.1 coding sequence.



XX red fluorescent protein; DsRed; fluorescence; red wavelength;  
KM oligomerization; tetramerization; immunoassay; hybridization assay; gene;  
KW ds.  
XX Discosoma sp.  
OS  
XX Key Location/Qualifiers  
FH CDS 1..678  
FT /\*tag= a  
FT /product= "DsRed variant mRFP1.1 protein"  
XX WO2003086446-A1.  
PN  
XX 23-OCT-2003.  
PD  
XX 09-APR-2003; 2003WO-US010879.  
PF  
XX 10-APR-2002; 2002US-00121258.  
PR 29-JUL-2002; 2002US-00209208.  
PR 29-JUL-2002; 2002US-00209208.  
XX (REGC ) UNIV CALIFORNIA.  
PA  
XX Tsien RY, Campbell RE, Baird GS;  
PI  
XX WPI; 2003-845265/78.  
DR P-PSDB; ADL46281.  
XX  
XX New monomeric and dimeric Anthozoan fluorescent protein variants with  
PT reduced propensity to oligomerize, and encoding polynucleotides, useful  
PT in molecular biology, e.g. in immunoassays or in tracking protein  
PT movement in cells.  
XX  
XX Claim 63, SEQ ID NO 80; 166pp; English.  
PS  
XX  
XX The invention relates to a polynucleotide sequence encoding a Discosoma  
CC red fluorescent protein (DsRed) variant having a reduced propensity to  
CC oligomerize. The protein variant comprises one or more amino acid  
CC substitutions at the AB and/or AC interface(s) of the wild-type DsRed  
CC sequence, where the substitutions result in reduced propensity of the  
CC DsRed variant to form tetramers and where the variant displays detectable  
CC fluorescence of at least one red wavelength. The composition and methods  
CC are useful in producing red fluorescent proteins having reduced  
CC propensity for oligomerization, especially tetramerization. The protein  
CC may be used in molecular biology and in other scientific applications,  
CC such as in immunoassays or hybridization assays, or in tracking the  
CC movement of proteins in cells. This sequence corresponds to the DsRed  
CC variant mRFP1.1 coding sequence. The protein contains a Q66M and a T147S  
CC mutation.  
XX  
XX Sequence 678 BP; 148 A; 227 C; 208 G; 95 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 2.45e-115 Length: 678  
Score: 1012.00 Matches: 190  
Percent Similarity: 88.69% Conservative: 6  
Best Local Similarity: 85.97% Mismatches: 25  
Query Match: 83.36% Indels: 0  
DB: 11 Gaps: 0

US-10-006-922A-12 (1-225) x ADL46282 (1-678)

QY 1 MetArgSerSerLySAAsnValIlleLySGluPheMetArgPheLyValArgMetGluGly 20  
||| |:::| |  
Db 1 ATGGCTCTCTCCGAGACGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 60  
  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
:::| |:::| |  
Db 61 TCCGTGAACGGCCACGAGTTGAGATGAGGGCGAGGGCGGCGCCCTACGAGGGC 120  
  
QY 41 HisAsnThrValysLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
||| |:::| |  
Db 121 ACCAGACCGCCAAAGTGAAGGTGACCAAGGGCGGCCCTGCTTGGCTGGACATC 180

QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrVallySHisProAlaAspIlePro 80  
||| |:::| |  
Db 181 CTGTCCCTCAGTTCATGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGCGACATCCCC 240  
  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
||| |:::| |  
Db 241 GACTACTTGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300  
  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
||| |:::| |  
Db 301 GACGGCGCGTGTGACCGTGAACCCAGACTCTCTCCAGAGACGGCGAGTTCAATCTAC 360  
  
QY 121 LysVallySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThr 140  
||| |:::| |  
Db 361 AAGGTGAAGCTGCGCGGACCAACTTCCCTCCGACGGCCCCGTATGACAGAGAGACC 420  
  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySlyGlu 160  
||| |:::| |  
Db 421 ATGGGCTGGAGGCTCTCTCCGAGCGGATGTATCCCGAGAGACGGCGCTGAAGGGCGAG 480  
  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180  
||| |:::| |  
Db 481 ATCAAGATGAGGTGAAGCTGAAGACGCGGCCACTACGACGCCGAGGTCAAGACACC 540  
  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
||| |:::| |  
Db 541 TACATGGCCAAAGAGCCCGTGACGCTGCGCGCCCTACAAAGACGACATCAAGCTTGAC 600  
  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
||| |:::| |  
Db 601 ATCACTCCCAACAAGAGACTACACCATCTGTGAACAGTACGAGCGCGCGAGGCCCGC 660  
  
QY 221 His 221  
|||  
Db 661 CAC 663

Search completed: January 12, 2006, 12:55:37  
Job time : 528 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:26:48 ; Search time 3746 Seconds  
(without alignments)  
2810.223 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIVEQYERTEGRHLEFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US10006922/runat\_10012006\_162504\_19128/app\_query.fasta\_1.391  
-DB=EST\_QFMT=fastcap -SUFFIX=p2nmlnscr.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=69 -ALIGN=500 -MODE=LOCAL  
-OUTFMT=plco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922\_QCGN\_1\_1\_8010 @runat\_10012006\_162504\_19128 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hlc:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_gss1:  
10: gb\_gss2:  
11: gb\_gss3:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
No matches found					

Search completed: January 12, 2006, 14:59:07  
Job time : 3747 secs



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:40:13 ; Search time 179 Seconds  
(without alignments)  
2234.365 Million cell updates/sec

Title: US-10-006-922A-12

Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTIIVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.epool\_p/US10006922/runat.10012006.162505.19142/app\_query.fasta\_1.391  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2nm1nscr.rn1 -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=500 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=69 -ALIGN=500 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500  
-MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CCG 1.1 290 @runat.10012006.162505.19142 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1214	100.0	678	US-09-459-956-6	Sequence 6, Appli
2	1214	100.0	859	US-09-866-538-11	Sequence 11, Appl
3	1214	100.0	859	US-09-865-291-11	Sequence 11, Appl
4	1210	99.7	723	US-10-152-296-1	Sequence 1, Appli
5	1210	99.7	6984	US-10-001-189-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1  
US-09-459-956-6  
; Sequence 6, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp  
US-09-459-956-6

Alignment Scores:  
Pred. No.: 3.24e-160 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 3

US-10-006-922A-12 (1-225) x US-09-459-956-6 (1-678)

QY	1	MetArgSerSerLyAsnValIleValGluPheMetArgPheLyValArgMetGluGly	20
DB	1	ATGAGGCTCTTCCAAGATGTTATCAAGAGTTCATGAGGTTAAGTTGCGATGGAAGGA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACGGTCAATGGGACAGAGTTGAATAGAGCGAAGAGAGGGGAGCCATACGAAGGC	120
QY	41	HisAsnThrValLySLeuLyValThrLySGLyGlyProLeuProPheAlaTrpAspIle	60
DB	121	CACAATACCGTAAGCTTAAGGTAACCAAGGGGACCTTGCCATTGCTGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro	80
DB	181	TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTCCGACATACCA	240
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrPGLuArgValMetAsnPheGlu	100
DB	241	GACTATAAAAGCTGTCATTCTCTGAAGATTAAATGGAAAGGGTCATGAACCTTGAA	300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr	120
DB	301	GACGGTGGCGTCTACTGTAACCAAGATTCAGTTTGACAGATGGCTGTTCATCTAC	360
QY	121	LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr	140
DB	361	AAGGTCAAGTTCTTGGCGTGAACCTTCCTCGATGACCTGTATGCAAAAGAAGACA	420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu	160
DB	421	ATGGGCTGGGAAGCAGCACTGAGCGTTTGTATCTCTGATGGCGTGTGAAGAGAG	480
QY	161	IleHisLySAlaLeuLySLeuLyBAspGlyGlyHisTyrLeuValGluPheLySLeu	180
DB	481	ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAATATT	540



```
QY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp   200
        |||
DB       541 TACATGGCAAGAAGCCTGTGCAGTACCAGGTA CCGTACTATATTGTTGATCCAACTTGAT   600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrgluArgThrglnGlyArg   220
        |||
DB       601 ATAACAAGCCACAACGAAGACTATACTCGTTGAGCAGTATGAAGAAGAACGAGGAGCACGC   660

QY      221 HisHisLeuPheLeu   225
        |||
DB       661 CACCATCTGTT CCTT   675
```

```

RESULT 2
US-09-866-538-11
; Sequence 11, Application US/09866538
; Patent No. 6852849
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEH, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-866-538-11

```

Alignment Scores:	
Pred. No.:	4.72e-160
Score:	1214.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	3
Length:	859
Matches:	225
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-09-866-538-11 (1-859)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	54	ATGAGGCTTCCAGAAGATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCCATGGAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyIleArgProTyrGluGly	40
Db	114	ACGGTCAATGGGCACGAGTTTGAATAGAAGCCGAAGAGAGAGGGAGGCCATACGAAGGC	173
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	174	CACATACCGTAAAGCTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	234	TTGTCAACCACAAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCCCTGCCACATACCA	293
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrPgluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAAGCTGTCAATTCCTGAAGATTAAATGGGAAAGGGTCATGAACTTGAA	353
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	354	GACGGTGGCGTCGTTACTGTAAACCGAGATTCACGTTTGCAAGATGGCTGTTTCATCTAC	413
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	414	AAGGTCAAGTTCAATGGCGGTGAACATTCCTCCGATGGAACCTGTTATGCAAAAGAAACA	473

QY	141	MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	474	ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTGATGGCGGTGTGAAGAAGAGAG	533
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	534	ATTCATAAGGCTCTGAAGCTGAAGAAGACGGTGTGCATTACTAGTTGAATTCAAAAGTATT	593
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValaspSerLysLeuAsp	200
Db	594	TACATGGCAAGAAGACCTGTGCAGCTACCAAGGTACTACTACTGTGACTCCAAACTGGAT	653
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	654	ATAACAAGCCACACGACGACTATACAATCGTTGAGCAGTATGAAGAACAACCGAGGAGCGC	713
QY	221	HisHisLeuPheLeu	225
Db	714	CACCATCTGTTCCCTT	728

```

RESULT 3
US-09-865-291-11
; Sequence 11, Application US/09865291
; Patent No. 6900304
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEH, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RADIOMETRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Drosophila sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54) .. (731)
US-09-865-291-11

```

Alignment Scores:	
Pred. No.:	4.72e-160
Score:	1214.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	3
Length:	859
Matches:	225
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-09-865-291-11 (1-859)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	54	ATGAGGCTTCCAAAGAATGTTATCAAGAGGTTCATGAGGTTTAAAGTTTCGATGGAAGGA	113
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	114	ACGGTCATATGGGCACGAGTTTGAATAAGAAGCGAAGAGAGGGGAGCCATACGAAGGC	173
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	174	CACATATACCGTAAGCTTAAAGTAACCAAGGGGGGACCTTGGCAATTGGCTGGGATATT	233
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	234	TTGTCAACCACAATTTCAGTATGAGAACCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	293
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	294	GACTATAAAAGCTGTCTATTCTCTGAAAGATTTAATGGGAAGGTCATGAACCTTTGAA	353



QY 101 AspglyglyvalvalthrvalthrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||  
DB 354 GACGGTGGCGTCTGTAAGTAAAGCCAGGATTCAGTTGCGAGATGGCTGTTCACTAC 413  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
|||  
DB 414 AAGGTCAAGTTCAATGGCGGTGAAGCTTTCTTCCTCCGATGACCTGTATGCAAAAGAGACA 473  
QY 141 MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyIle 160  
|||  
DB 474 ATGGGCTGGGAAGCCAGCAGTGAAGCGGTTGTATCCGTGATGGCGGTGTAAGAGAG 533  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||  
DB 534 ATTCATAAGGCTCTGAAGCTGAAGACGGTGTCTATACCTAGTTGAATTCAAAAGTATT 593  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||  
DB 594 TACATGGCAAGAAGCCTGTGACGTAACCAAGGTAATACTGTAAGTCAAACTGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||  
DB 654 ATAAACAAGCCACAACGAAAGACTATACATCGTTGAGCAGTATGAAGAACCGAGGAGCG 713  
QY 221 HisHisLeuPheLeu 225  
|||  
DB 714 CACCAATCTGTTCCCTT 728

## RESULT 4

US-10-152-296-1  
; Sequence 1, Application US/10152296  
; Patent No. 6723537  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/152,296  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 723  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DsRED) of Drososoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(723)  
; OTHER INFORMATION: DsRED  
US-10-152-296-1

## Alignment Scores:

Pred. No.:	1.31e-159	Length: 723
Score:	1210.00	Matches: 224
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.56%	Mismatches: 0
Query Match:	99.67%	Indels: 0
DB:	3	Gaps: 0

US-10-006-922A-12 (1-225) x US-10-152-296-1 (1-723)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
:::|||||  
DB 4 GTGCGCTCCTCAAGACGTCAATCAAGAGAGTCAATGCGCTTCAAGGTGCGCATGGAAGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40

DB 64 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCCGAGGGCCGCCCTCAAGAGGC 123  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
|||  
DB 124 CACAACAACCGTGAAGCTGAAGGTGAAGCAAGGGCGGCCCTGCCCCCTTCCGCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
|||  
DB 184 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGTACGTGAAGACACCCCGCGACATCCCC 243  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
|||  
DB 244 GACTACAAGAAGCTGTCTTCCCGGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 303  
QY 101 AspglyglyvalvalthrvalthrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
|||  
DB 304 GACGGCGCGGTGTGACCGTGAACCAAGACTCCTCCCTGCAAGACGGCTGCTCATCTAC 363  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIleThr 140  
|||  
DB 364 AAGGTGAAGTTCAATCGCGGTGAAGCTTCCCTCCGACGGCCCGTATGCAAGAAAGACC 423  
QY 141 MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyIle 160  
|||  
DB 424 ATGGGCTGGGAAGCCTCCACGAGCGCCTGTACACCCCGCGAGCGCGTGTGAAGGGCGAG 483  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
|||  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGGCCACTACTGTTGAAGTTCAAGAGTATC 543  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
|||  
DB 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGAAGTCAAGCTGAGC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
|||  
DB 604 ATCACTCCCAACAAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 663  
QY 221 HisHisLeuPheLeu 225  
|||  
DB 664 CACCACTGTTCCTG 678

## RESULT 5

US-10-001-189-45  
; Sequence 45, Application US/10001189  
; Patent No. 6962810  
; GENERAL INFORMATION:  
; APPLICANT: FRASER JR., MALCOLM J.  
; APPLICANT: LI, XU  
; APPLICANT: BEAM, TERESA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
; FILE REFERENCE: 835910-92098  
; CURRENT APPLICATION NUMBER: US/10/001,189  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/244,984  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/244,677  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 45  
; LENGTH: 6984  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: pBXF3-DsRed-ori sequence  
US-10-001-189-45

## Alignment Scores:

Pred. No.:	1.31e-159	Length: 723
Score:	1210.00	Matches: 224
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	99.56%	Mismatches: 0
Query Match:	99.67%	Indels: 0
DB:	3	Gaps: 0



Pred. No.:	4.87e-158	Length:	6984
Score:	1210.00	Matches:	224
Percent Simlarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	3	Gaps:	0

US-10-006-922A-12 (1-225) x US-10-001-189-45 (1-6984)

QY	1	MetArgSerSerLyAsnValIleYsGluPheMetArgPheLyValArgMetGluGly	20
Db	2974	GTGGCTCTCCCAAGAACTCAACAGAGTTTCATGCGCTTCAAGTGCGCATGAGGCG	3033
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	3034	ACCGTGAACGGCCACGAGTTCGAGTTCGAGGGCGAGGGCGCCCTACGAGGCG	3093
QY	41	HisAsnThrValLyLeuLyValThrLySgLyGlyProLeuProPheAlaTrpAspIle	60
Db	3094	CACAAACCCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGGACATC	3153
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro	80
Db	3154	CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGACGTGAACACCCCGGACATCCCC	3213
QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
Db	3214	GACTAACAAGACTGTCTCTCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG	3273
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	3274	GACGGCGGCGTGTGACCGTGACCAAGACTCTCCCTGCGAGACGGCTGTTCATCTAC	3333
QY	121	LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeThr	140
Db	3334	AAGGTGAAGTTCATCGCGGTGAACCTCCCTCCGACGGCCCGTAAAGCAAGAAAGACC	3393
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySgLyGlu	160
Db	3394	ATGGGCTGGGAGGCTCCACCGAGCGCCTGTACCCCGCGACGGCGGTGAAAGGGCGAG	3453
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeIle	180
Db	3454	ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCACACTACTACTACTGAGTTCAAAGTTCATC	3513
QY	181	TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
Db	3514	TACATGGCCAAAGAGCCCGTGCACTGCCCGCTACTACTACTGAGTCCCAAGCTGGAC	3573
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	3574	ATCACTCCCAACAAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCGCGC	3633
QY	221	HisHisLeuPheLeu	225
Db	3634	CACCACTGTTCCTG	3648

Search completed: January 12, 2006, 15:02:20  
 Job time : 182 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:46:57 ; Search time 792 Seconds  
(without alignments)  
2349.255 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTYVEQYERTEGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 95

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/cg2\_1/USPTO.spool.p/US10006922/runat.10012006.162505.19171/app\_query.fasta\_1.391  
-DB=Published\_Applications\_NA\_Main -QFMT=Fastap -SUFFIX=p2nmInscr.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=500 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=69 -ALIGN=500 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10006922 @CG2\_1.1549 @runat.10012006.162505.19171 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT\_DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA\_Main:\*

- 1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cg2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cg2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cg2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cg2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cg2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cg2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cg2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cg2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1214	100.0	678	3	US-09-967-772-6 Sequence 6, Appli
2	1214	100.0	678	5	US-10-006-922-11 Sequence 11, Appli
3	1214	100.0	678	5	US-10-081-864-7 Sequence 7, Appli
4	1214	100.0	678	5	US-10-121-258-2 Sequence 2, Appli
5	1214	100.0	678	6	US-10-315-920-1 Sequence 1, Appli
6	1214	100.0	678	6	US-10-132-067-3 Sequence 3, Appli
7	1214	100.0	678	6	US-10-335-517-6 Sequence 6, Appli

8	1214	100.0	678	6	US-10-334-288-6 Sequence 6, Appli
9	1214	100.0	678	7	US-10-311-030-5 Sequence 5, Appli
10	1214	100.0	678	8	US-10-656-029-21 Sequence 21, Appli
11	1214	100.0	678	9	US-10-505-486-27 Sequence 27, Appli
12	1214	100.0	678	9	US-10-844-064A-1 Sequence 1, Appli
13	1214	100.0	678	9	US-10-931-304-2 Sequence 2, Appli
14	1214	100.0	859	3	US-09-999-745-66 Sequence 66, Appli
15	1214	100.0	859	3	US-09-866-538-11 Sequence 11, Appli
16	1214	100.0	859	3	US-09-794-308-11 Sequence 11, Appli
17	1214	100.0	859	3	US-09-865-291-11 Sequence 12, Appli
18	1214	100.0	859	7	US-10-433-640-12 Sequence 11, Appli
19	1214	100.0	859	8	US-10-885-988-11 Sequence 11, Appli
20	1214	100.0	859	3	US-10-857-622-11 Sequence 11, Appli
21	1214	100.0	3311	3	US-09-797-496B-3 Sequence 3, Appli
22	1210	99.7	681	5	US-10-121-258-3 Sequence 3, Appli
23	1210	99.7	681	5	US-10-311-030-8 Sequence 8, Appli
24	1210	99.7	681	5	US-10-311-030-8 Sequence 3, Appli
25	1210	99.7	681	9	US-10-931-304-3 Sequence 23, Appli
26	1210	99.7	681	9	US-10-931-304-23 Sequence 11, Appli
27	1210	99.7	713	7	US-10-311-030-11 Sequence 12, Appli
28	1210	99.7	713	7	US-10-152-296-1 Sequence 1, Appli
29	1210	99.7	723	5	US-10-739-656-1 Sequence 1, Appli
30	1210	99.7	1638	5	US-10-214-932-51 Sequence 51, Appli
31	1210	99.7	1647	5	US-10-214-932-75 Sequence 75, Appli
32	1210	99.7	4692	6	US-10-161-403-29 Sequence 29, Appli
33	1210	99.7	4692	7	US-10-433-640-16 Sequence 16, Appli
34	1210	99.7	4692	10	US-11-006-076-29 Sequence 29, Appli
35	1210	99.7	5436	9	US-10-169-050-46 Sequence 46, Appli
36	1210	99.7	5436	9	US-10-001-189-45 Sequence 45, Appli
37	1210	99.7	6984	5	US-10-169-050-20 Sequence 20, Appli
38	1210	99.7	7910	9	US-10-471-065-20 Sequence 4, Appli
39	1210	99.7	9320	8	US-10-609-019-4 Sequence 4, Appli
40	1210	99.7	9658	8	US-10-609-019-3 Sequence 3, Appli
41	1210	99.7	9678	8	US-10-609-019-2 Sequence 2, Appli
42	1210	99.7	10263	6	US-10-315-920-3 Sequence 3, Appli
43	1210	99.4	678	7	US-10-742-828-4 Sequence 4, Appli
44	1207	99.4	7508	7	US-10-006-922-37 Sequence 37, Appli
45	1203	99.1	681	5	US-10-314-936-1 Sequence 1, Appli
46	1198	98.7	666	7	US-10-314-936-3 Sequence 3, Appli
47	1196	98.5	711	7	US-11-021-014-1 Sequence 1, Appli
48	1196	98.5	711	7	US-11-021-014-1 Sequence 3, Appli
49	1196	98.5	711	10	US-10-081-864-14 Sequence 14, Appli
50	1196	98.5	711	10	US-10-315-920-5 Sequence 5, Appli
51	1196	98.5	711	10	US-10-006-922-38 Sequence 38, Appli
52	1191	98.1	678	5	US-10-081-864-13 Sequence 13, Appli
53	1191	98.1	678	5	US-10-081-864-13 Sequence 10, Appli
54	1191	98.1	678	6	US-10-060-922-45 Sequence 45, Appli
55	1186	97.7	675	5	US-10-121-258-5 Sequence 5, Appli
56	1186	97.7	675	5	US-10-931-304-5 Sequence 5, Appli
57	1186	97.7	747	8	US-10-844-064A-3 Sequence 40, Appli
58	1186	97.7	747	8	US-10-423-688A-40 Sequence 7, Appli
59	1186	97.7	7495	9	US-10-121-258-7 Sequence 7, Appli
60	1186	97.7	4300	9	US-11-052-001-5 Sequence 5, Appli
61	1186	97.7	4300	9	US-10-006-922-43 Sequence 43, Appli
62	1186	97.7	704	9	US-10-081-864-11 Sequence 11, Appli
63	1186	97.7	705	9	US-10-081-864-11 Sequence 82, Appli
64	1184	97.5	1431	9	US-10-931-304-107 Sequence 107, Appli
65	1160	95.6	876	5	US-10-006-922-17 Sequence 17, Appli
66	1160	95.6	876	6	US-10-161-403-39 Sequence 39, Appli
67	1160	95.6	876	10	US-11-006-076-39 Sequence 39, Appli
68	1132	93.2	678	7	US-10-724-178-15 Sequence 15, Appli
69	1121	92.3	681	5	US-10-121-258-9 Sequence 9, Appli
70	1121	92.3	681	9	US-10-931-304-7 Sequence 7, Appli
71	1121	92.3	1809	10	US-10-006-922-43 Sequence 43, Appli
72	1119	92.2	678	5	US-10-081-864-11 Sequence 11, Appli
73	1119	92.2	705	9	US-10-931-304-82 Sequence 82, Appli
74	1095	90.2	1431	9	US-10-931-304-107 Sequence 107, Appli
75	1095	90.2	876	5	US-10-006-922-17 Sequence 17, Appli
76	1085.5	89.4	876	6	US-10-161-403-39 Sequence 39, Appli
77	1085.5	89.4	876	10	US-11-006-076-39 Sequence 39, Appli
78	1085.5	89.4	876	10	US-10-724-178-15 Sequence 15, Appli
79	1021	84.1	675	7	US-10-121-258-9 Sequence 9, Appli
80	1021	84.1	678	5	US-10-931-304-7 Sequence 7, Appli



81	1021	84.1	678	9	US-10-931-304-9	Sequence 9, Appli
82	1014	83.5	678	9	US-10-931-304-109	Sequence 109, App
83	1012	83.4	678	9	US-10-931-304-80	Sequence 80, Appl
84	1007	82.9	678	9	US-10-931-304-105	Sequence 105, App
85	1002	82.5	678	9	US-10-931-304-101	Sequence 101, App
86	994	81.9	711	9	US-10-931-304-86	Sequence 86, Appl
87	992	81.7	711	9	US-10-931-304-90	Sequence 90, Appl
88	989	81.5	711	9	US-10-931-304-103	Sequence 103, App
89	988	81.4	711	9	US-10-931-304-84	Sequence 84, Appl
90	987	81.3	711	9	US-10-931-304-93	Sequence 93, Appl
91	986	81.2	708	9	US-10-931-304-95	Sequence 95, Appl
92	973	80.1	711	9	US-10-931-304-97	Sequence 97, Appl
93	973	80.1	711	9	US-10-931-304-99	Sequence 99, Appl
94	960	79.1	711	9	US-10-931-304-88	Sequence 88, Appl
95	848	69.9	549	7	US-10-724-178-1040	Sequence 1040, Ap

ALIGNMENTS

RESULT 1  
US-09-967-772-6  
; Sequence 6, Application US/09967772  
; Patent No. US20020164577A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/09/967,772  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp "red"  
US-09-967-772-6

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922a-12 (1-225) x US-09-967-772-6 (1-678)

QY	1	MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly	20
DB	1	ATGAGGCTCTCCAAAGATGTATCAAGAGGTTTCATGAGGTTTAAAGTTCGCAATGAGGA	60
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
DB	61	ACGGTCAATGGGACGAGTTTGAATAAGAAGCGAAGGAGGAGGAGCCATACGAGGC	120
QY	41	HisAsnThrVallySLeuLySValThrLySGlyGlyProLeuProPheAlaTrrAspIle	60
DB	121	CACATACCGTAAAGCTTAAGGTAAACCAAGGGGGGACCTTGCCATTGGCTTGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrGlySerLySValTyrVallyHisProAlaAspIlePro	80
DB	181	TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCA	240

QY	81	AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu	100
DB	241	GACTATAAAAAGCTGTCAATTCCTCGAAGGATTTAAATGGAAAGGTCATGAACTTGAA	300
QY	101	AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
DB	301	GACGGTGGCGTGTACTGTAAACCCAGATTCAGTTTGACAGATGGCTGTTTCATCTAC	360
QY	121	LySVallySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr	140
DB	361	AAGTCAAGTTCATTGGCGTGAACCTTCCTCCGATGACCTGTTATGCAAAAGAGACA	420
QY	141	MetGlyTrrPgluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlu	160
DB	421	ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCTCTCGTATGGCGGTGTGAAGAGAG	480
QY	161	IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu	180
DB	481	ATTCAATAGGCTGTGAAGCTGAAAGACGGTGTCAATTACCTAGTTGAATTCAAAATATT	540
QY	181	TyrMetAlaLySlySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp	200
DB	541	TACATGCCAAAGAGCCTGTGCAGCTACCAAGGTACTATGTTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluArg	220
DB	601	ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGGAGCGC	660
QY	221	HisHisLeuPheLeu	225
DB	661	CACCATCTGTCTCTT	675

RESULT 2  
US-10-006-922-11  
; Sequence 11, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1e1 Chromophores/Fluorophores and  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/457,556  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/444,338  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma species  
US-10-006-922-11

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0



```

DB:          5          Gaps:          0

US-10-006-922a-12 (1-225) x US-10-006-922-11 (1-678)

QY      1 MetArgSerSerLyAsnValIleIysGluPheMetArgPheLySValArgMetGluGly 20
        |||||||
Db      1 ATGAGGCTCTCCAAGAATGTATCAAGAGTTTCATGAGGTTTAAGTTTCGATGAAGGA 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
        |||||||
Db      61 ACCGTCAATGGGCACGAGTTGAATAAGAAGCGAAGAGAGGAGGCCATACGAAGGC 120

QY      41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTpaAspIle 60
        |||||||
Db      121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCCTTGGCAATTGCTGGGATATT 180

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisPProAlaAspIlePro 80
        |||||||
Db      181 TTGTCAACCACAATTTCAGTATGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 240

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
        |||||||
Db      241 GACTATATAAAAGCTGTTCATTTCTGAAAGATTAAATGGAAAGGTCATGAACTTTGAA 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        |||||||
Db      301 GACGGTGGCGCTGTTACTGTAAACCCAGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC 360

QY      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
        |||||||
Db      361 AAGGTCAAGTTTCATTGGCGTGAACCTTCTTCGATGGAACCTGTTATGCAAAAGAAGACA 420

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160
        |||||||
Db      421 ATGGGCTGGGAAGCCAGCACTGAGCGGTTGTATCTCTGATGGCGGTTTGAAAAGAGAG 480

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180
        |||||||
Db      481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGCATTACCTAGTGAATTCAAAAGTATT 540

QY      181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
        |||||||
Db      541 TACATGGCAAGAAGCCTGTGACGTAACGAGGTACTACTATGTGATGCCAAACTGGAT 600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
        |||||||
Db      601 ATAACAAGCCCAACAAGACTATACATCGTTGACAGTATGAAGAACCGAGGAGCGC 660

QY      221 HisHisLeuPheLeu 225
        |||||||
Db      661 CACCATCTGTCTCTT 675

RESULT 3
US-10-081-864-7
; Sequence 7, Application US/10081864
; Publication No. US20030022287A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey
; APPLICANT: Lukyanov, Konstantin
; APPLICANT: Yanushevich, Yuriy
; APPLICANT: Savitsky, Alexandr
; APPLICANT: Fradkov, Arcady
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-067
; CURRENT APPLICATION NUMBER: US/10/081,864
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: 10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/270,983
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 7
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp
US-10-081-864-7

Alignment Scores:
Pred. No.:          6.83e-143          Length:          678
Score:             1214.00             Matches:          225
Percent Similarity: 100.00%             Conservative:      0
Best Local Similarity: 100.00%             Mismatches:       0
Query Match:       100.00%             Indels:           0
DB:                5                   Gaps:              0

US-10-006-922a-12 (1-225) x US-10-081-864-7 (1-678)

QY      1 MetArgSerSerLyAsnValIleIysGluPheMetArgPheLySValArgMetGluGly 20
        |||||||
Db      1 ATGCGCTCCTCCAAGAAGCTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
        |||||||
Db      61 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCCCTACGAGGGC 120

QY      41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTpaAspIle 60
        |||||||
Db      121 CACAACAACGTGAAGCTGAAGGTGAACCAAGGGCGGCCCTTGCCCTTGCGCTGGGACATC 180

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisPProAlaAspIlePro 80
        |||||||
Db      181 CTGTCCCCCAAGTTCAGATACGGCTTCCAAGGTGTACGTGAAGCAACCCGCGACATCCCC 240

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
        |||||||
Db      241 GACTACAAGAAGCTGTCTCTTCCCCAGGGGCTTCAAGTGGAGCGCGTGAATGAACCTTGAG 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
        |||||||
Db      301 GACGGCGGCGTGTGACCGGTGACCCAGACTCTCCCTGCAAGAGGCTGCTTCATCTAC 360

QY      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
        |||||||
Db      361 AAGGTGAAGTTTCATCGCGTGAACCTTCCCCCTCCGACGGCCCCGTGATGCAAGAAGACC 420

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160
        |||||||
Db      421 ATGGGCTGGAGGCTTCCACCGAGCGGCTGTAAACCCCGCAGCGGCGTGTGAAGGCGAG 480

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180
        |||||||
Db      481 ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGCACTACCTGTGTGAGTTCAAGTCCATC 540

QY      181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
        |||||||
Db      541 TACATGGCCAAGAAGCCGTGACGTCGCCGCTACTACTACGTGAGTCCAAGCTGGAC 600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
        |||||||
Db      601 ATCACCCTCCACAACGAGACTACACCATCGTGTGAGAGTACGAGCGCACCGAGGGCGCG 660

QY      221 HisHisLeuPheLeu 225
        |||||||
Db      661 CACCACCTGTCTCTG 675

RESULT 4
US-10-121-258-2
; Sequence 2, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

```



FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/10/121,258  
CURRENT FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(678)  
OTHER INFORMATION: wild-type DeRed  
US-10-121-258-2

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-2 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGAGGTCTTCCAAGAATGTTATCAAGAGTTCATGAGGTTAAGTTCGCATGAGAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTCAATGGGCACGAGTTGAAATGAAGGCGAAGGAGGGGAGGCCATACGAAGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTpaApIle 60  
Db 121 CACAATAACGTAAGCTTAAGGTAAACCAAGGGGGACCTTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTCAGTATGAGCAAGATATATGCAAGCACCTGCCGACATACCA 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAAAGCTGTCATTCTCGAAGATTAAATGGGAAAGGTCATGAACCTTGA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCGTACTGTATCAACAGATTCCAGTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTCAAGTTCATTGGCGCTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTrrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACGTGAGCGTTGTATCCTCGTATGGCGGTGTAAGGAGAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
Db 481 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTATGTTGAATTCAAAAGTATT 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGCTACTACTATGTTGACTCCAAACTGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCAGGGGACGC 660

QY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCCTT 675

RESULT 5  
US-10-315-920-1  
Sequence 1, Application US/10315920  
Publication No. US20030175809A1  
GENERAL INFORMATION:  
APPLICANT: Fradkov, Arcady Fedorovich  
APPLICANT: Tersikh, Alexey  
TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
TITLE OF INVENTION: FOR THEIR USE  
FILE REFERENCE: CLON-077CIP  
CURRENT APPLICATION NUMBER: US/10/315,920  
CURRENT FILING DATE: 2002-12-09  
PRIOR APPLICATION NUMBER: 60/211,607  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: PCT/US01/19097  
PRIOR FILING DATE: 2001-06-13  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: variant of sequence from Discosoma sp.  
US-10-315-920-1

Alignment Scores:  
Pred. No.: 6,83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-315-920-1 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
Db 1 ATGCGCTCCTCCAAGACGTATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTpaApIle 60  
Db 121 CACAACACCGTAAGCTGAAGGTGAACCAAGGGCGGCCCTTCCCTTCCGCTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTGTACGTGAAGCACCCCGGACATCCC 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAAGCGGTGATGAACCTTGA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGTGTGACCGTGAACCAAGACTCTCCTGCAAGACGGCTGCTCATCTAC 360  
QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTGATGCAAGAAGACC 420  
QY 141 MetGlyTrrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCTCAACGAAGCGCCTGTACCCCCCGGACGGCGTGTGAAGGGCAG 480  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180



Db	481	ATCCACAAGCCCTGAGCTGAGAGACGCGGGCACTACCTGGTGAGTTCAAGTCCATC	540
QY	181	TYRMetAlaIysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGCCCCAAGAAGCCCGTCAGCTGCCGCTACTACTAGTGAGCTCCAGCTGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGlnIleArg	220
Db	601	ATCACTCCCAACAAGAGACTACCACTCTGTGAGCAGTACGAGCGCACCGAGGCGCC	660
QY	221	HisHisLeuPheLeu	225
Db	661	CACCACCTGTTCCTG	675

```

RESULT 6
US-10-132-067-3
; Sequence 3, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands with Intrinsic
; TITLE OF INVENTION: Fluorescence
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discozyma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: dsRED
;
US-10-132-067-3

```

Alignment Scores:	6.83e-143	Length:	678
Pred. No.:	1214.00	Matches:	225
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-006-922A-12 (1-225) X US-10-132-067-3 (1-678)

QY	1	Me	Arg	Ser	Ser	Le	Val	Val	I	Le	Val	S	Le	u	P	he	Met	Arg	P	he	Val	Arg	Met	G	Le	u	20													
Db	1	AT	GAG	GCT	TT	CC	AA	GA	TG	TAT	CA	AG	GAG	TT	CA	TG	AG	GTT	TA	AG	GTT	CC	GA	TG	GA	AG	60													
QY	21	Thr	Val	Asn	G	L	Y	H	I	S	G	L	P	he	G	L	I	Le	G	L	I	Y	G	L	I	Y	Arg	Pro	Ty	G	Le	u	40							
Db	61	AC	GGT	CA	AT	GG	CA	CG	AG	TT	TG	AA	AT	AG	AA	GG	CG	AA	GA	GG	GG	AG	GC	CA	TA	CG	AA	GC	120											
QY	41	H	is	Asn	Thr	Val	Le	u	Le	u	Le	u	Val	Thr	Le	u	S	G	L	I	Y	G	L	I	Y	Pro	Leu	Pro	Ph	ea	I	at	Tr	Asp	I	le	60			
Db	121	CA	CA	AT	AC	CG	T	AA	A	G	CT	T	AA	G	TA	AC	CA	A	G	G	G	G	GA	C	TT	T	G	CC	AT	T	T	G	G	GA	T	AT	180			
QY	61	Leu	Ser	Pro	G	ln	P	he	G	ln	Ty	r	G	L	I	Y	Ser	Le	u	Val	Ty	r	Val	Le	u	His	Pro	Ala	Asp	I	le	Pro	80							
Db	181	TT	GT	CA	CC	CA	CA	AT	TT	CA	GT	AT	TG	GA	AG	CA	AG	GT	AT	AT	GT	CA	AG	CA	CC	CT	GC	CG	CA	TA	CC	240								
QY	81	Asp	Ty	r	Le	u	Le	u	Le	u	Ser	P	he	Pro	G	ln	G	L	I	Y	P	he	Le	u	S	Ty	r	P	G	L	u	Arg	Val	Met	Asn	P	he	G	ln	100
Db	241	GA	CT	AT	AAAA	AG	CT	GT	CA	TT	T	CT	GA	AG	GA	TT	T	AA	A	T	GG	GA	AA	AG	GT	CA	TG	AA	CT	T	T	GA	300							

QY	101	ASPG1GLYValValThrValThrglnaspSerSerLeuGlnaspGlyCyspheiIeTyr	120
DB	301	GACGGTGGCGTCTTACTGTAACCCAGGATTCAGTTTGAGGATGGCTGTTTCATCTAC	360
QY	121	LysValIlyspheiIegIYValasnpheroSeraspGlyProValmetGlnIlyslYsthr	140
DB	361	AAGSTCAAGTTCATGTGGCGTGAACTTCTCCGTGAGCACTGTTATGCAAAAGAAAGACA	420
QY	141	MetGlyTrpGluAlaSerThrgluArgleuTyrProArgaspGlyValleuIySGlyGlu	160
DB	421	ATGGGCTGGGAAGCCAGCACGTGACGCTTTGTATCCTCGTATGGCGCTGTGAAGAGAGAG	480
QY	161	IleHisIySAlaleuIyLeuIySaspGlyGlyHisIyTyrleuValGluPhelYsSerIle	180
DB	481	ATTCTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACTAGTGAATTCAAAAAGTATT	540
QY	181	TyrMetAlaIySlySProValGlnleuProGlyTyrTyrTyrValaspSerIySleuasp	200
DB	541	TACATGGCCAAAGAAGCCTGTGCAAGTACCAAGGTACTACTATGTTGACTCCAAACTGGAT	600
QY	201	IleThrSerHisasnGluaspTyrThrIleValGluGlnTyrGluArgThrgluGlyArg	220
DB	601	ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC	660
QY	221	HisHisIeuPhelIeu	225
DB	661	CACCATCTGTTCCTT	675

```

RESULT 7
US-10-335-517-6
; Sequence 6, Application US/10335517
; Publication No. US20030207248A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSIEN, Roger
; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/10/335,517
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US/09/967,772
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/ US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-10-335-517-6

```

Alignment Scores:	6.83e-143	Length:	678
Pred. No.:	1214.00	Matches:	225
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-006-922A-12 (1-225) x US-10-335-517-6 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGLuPhemetaArgPheLyValaArgMetGluLy 20  
DB 1 ATAGGCTCTCCAGAGATGTTATCAAGGAGTTCAAGGTTTAAGGTTCCCATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40



|||||  
Db 61 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCGCAAGGAGGGGAGGCCCATACGAAGCC 120  
Qy 41 HisAsnThrValIysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGGACCTTTGCCATTGTGGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTCAGTATGGAAGCAAGGATATATGCAAGCACCTGCGGACATACCA 240  
Qy 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAGCTGTCAATTCTCTGAAGGATTTAAATGGGAAAAGGTCATGAACCTTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTGCGAGGATGGCTGTTCACTTAC 360  
Qy 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCGATGACCTGTTATGCAAAAGAGACA 420  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGACGCTTTGTATCCTCGTGATGGCGTGTGAAGGAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTAAGTTGAATTCAAAAGTATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCACTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGACGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCCTT 675  
RESULT 8  
US-10-334-288-6  
; Sequence 6, Application US/10334288  
; Publication No. US20040002123A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: GONZALEZ, Jesus  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-5  
; CURRENT APPLICATION NUMBER: US/10/334,288  
; CURRENT FILING DATE: 2002-12-31  
; PRIOR APPLICATION NUMBER: US/09/967,772  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: US 08/765,860  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: PCT/ US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; PRIOR APPLICATION NUMBER: US 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discoosoma sp "red"  
US-10-334-288-6

Alignment Scores:  
Pred. No.: 6.83e-143 Length: 678  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-334-288-6 (1-678)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGGCTTCCAGAAATGTTATCAAGGAGTTCAATGAGTTTAAGTTCGCAATGAAGGA 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCAAGGTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAAGCTTAAAGGTAAACCAAGGGGGACCTTGCCATTGTGGGATATT 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValIysHisProAlaAspIlePro 80  
Db 181 TTGTCAACCAATTTCAGTATGGAAGCAAGGATATATGCAAGCACCTGCGGACATACCA 240  
Qy 81 AspTyrLysLysLeuSerPheProGlnGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAAGCTGTCAATTCTCTGAAGGATTTAAATGGGAAAAGGTCATGAACCTTGAA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTTACTGTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTCACTTAC 360  
Qy 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCGATGGAACCTGTTATGCAAAAGAGACA 420  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGACGCTTTGTATCCTCGTGATGGCGTGTGAAGGAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTAAGTTGAATTCAAAAGTATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCACTACCAAGGTACTACTATGTGACTCCAAACTGGAT 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAAGAACCGAGGACGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCCCTT 675  
RESULT 9  
US-10-311-030-5  
; Sequence 5, Application US/10311030  
; Publication No. US20040171107A1  
; GENERAL INFORMATION:  
; APPLICANT: Nelson, David  
; APPLICANT: Zamiatra, Elize  
; APPLICANT: Tsien, Roger  
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
; FILE REFERENCE: 15916-032US1  
; CURRENT APPLICATION NUMBER: US/10/311,030  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: PCT/US01/04625  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US 60/184,732



; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
US-10-311-030-5

Alignment Scores:
Pred. No.: 6,83e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-5 (1-678)

QY 1 MetArgSerSerIysAsnValIleIleYsGluPheMetArgPheIysValArgMetGluY 20
DB 1 ATGAGGCTTCCAGAAATGTTATCAAGAGATTCATGAGGTTTAAGGTCGCAATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 61 ACGGTCAATGGGCAAGAGTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC 120
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTyrAspIle 60
DB 121 CACAATACCGTAAGCTTAAGTAACCAAGGGGACCTTGGCATTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80
DB 181 TTGTCAACCAATTCAGTATGGAAGCAAGGATATGTCAAGCACCCCTGCCGACATACCA 240
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100
DB 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 301 GACGGTGGCGTCTGTAAGCTTAACCAAGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 360
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140
DB 361 AAGGTCAAGTTCATTTGGCGTGAACCTTCCCTCCGATGACCTGTTATGCAAAAGAGACA 420
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIu 160
DB 421 ATGGGCTGGGAAGCCAGCACTGACGTTGTATCCTCGATGCGGTGTGAAGGAGAG 480
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyIleIstYrLeuValGluPheIysSerIle 180
DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200
DB 541 TACATGGCAAAAGAGCCTGTGCACTACCAAGGTTACTATGTGACTCCAAACTGAT 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCGC 660
QY 221 HisHisIysLeuPheIu 225
DB 661 CACCATCTGTTCTT 675

RESULT 10
US-10-656-029-21
; Sequence 21, Application US/10656029
; Publication NO. US2005000367A1
; GENERAL INFORMATION:
; APPLICANT: VERTEX PHARMACEUTICALS INC.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF
; TITLE OF INVENTION: SCREENING ASSAYS
; FILE REFERENCE: VPI/02-143WO2
; CURRENT APPLICATION NUMBER: US/10/656,029
; PRIOR FILING DATE: 2003-09-05
; PRIOR APPLICATION NUMBER: 60/408,297
; PRIOR FILING DATE: 2002-09-05
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: fluorescent protein
US-10-656-029-21

Alignment Scores:
Pred. No.: 6,83e-143 Length: 678
Score: 1214.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-656-029-21 (1-678)

QY 1 MetArgSerSerIysAsnValIleIleYsGluPheMetArgPheIysValArgMetGluY 20
DB 1 ATGAGGCTTCCAGAAATGTTATCAAGAGATTCATGAGGTTTAAGGTTTCGCAATGGAAGGA 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
DB 61 ACGGTCAATGGGCAAGAGTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC 120
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTyrAspIle 60
DB 121 CACAATACCGTAAGCTTAAGTAACCAAGGGGACCTTGGCATTGCTTGGGATATT 180
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80
DB 181 TTGTCAACCAATTCAGTATGGAAGCAAGGATATGTCAAGCACCCCTGCCGACATACCA 240
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTyrGluArgValMetAsnPheGlu 100
DB 241 GACTATAAAAAGCTGTCAATTCCTGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 301 GACGGTGGCGTCTGTAAGCTTAACCAAGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 360
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140
DB 361 AAGGTCAAGTTCATTTGGCGTGAACCTTCCCTCCGATGACCTGTTATGCAAAAGAGACA 420
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIu 160
DB 421 ATGGGCTGGGAAGCCAGCACTGACGTTGTATCCTCGATGCGGTGTGAAGGAGAG 480
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyIleIstYrLeuValGluPheIysSerIle 180
DB 481 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 540
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200
DB 541 TACATGGCAAAAGAGCCTGTGCACTACCAAGGTTACTATGTGACTCCAAACTGAT 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCGC 660



QY 221 HisHisLeuPheLeu 225  
 Db 661 CACCATCTGTCTCCTT 675

RESULT 11

US-10-505-486-27  
 ; Sequence 27, Application US/10505486  
 ; Publication No. US20050118639A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Takeda Chemical Industries, Ltd.  
 ; TITLE OF INVENTION: Determination of a ligand  
 ; FILE REFERENCE: P03-0006PCT  
 ; CURRENT APPLICATION NUMBER: US/10/505,486  
 ; PRIORITY FILING DATE: 2004-08-20  
 ; PRIORITY APPLICATION NUMBER: JP 2002-45728  
 ; PRIORITY FILING DATE: 2002-02-22  
 ; PRIORITY APPLICATION NUMBER: JP 2002-213949  
 ; PRIORITY FILING DATE: 2002-07-23  
 ; PRIORITY APPLICATION NUMBER: JP 2002-298237  
 ; PRIORITY FILING DATE: 2002-10-11  
 ; NUMBER OF SEQ ID NOS: 233  
 ; SEQ ID NO 27  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: Discosoma sp.  
 US-10-505-486-27

Alignment Scores:  
 Pred. No.: 6.83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-505-486-27 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
 Db 1 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGTTTCGATGGAAGCA 60  
 QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGCAAGAGTTGAAATAGAACGCGAAGAGAGGGGAGGCCATACGAAGGC 120  
 QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAATACCGTAAGCTTAAGTAACCAAGGGGACCTTGGCATTTGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrgLySerLyValTyrrValLyHisProAlaAspIlePro 80  
 Db 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGTATATGTCAAGCACCTTCCGACATACCA 240  
 QY 81 AspTyrrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
 Db 241 GACTATAAAGCTGTCAATTTCTGGAAGATTAAATGGGAAAGGTCATGAACCTTGAA 300  
 QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrr 120  
 Db 301 GACGGTGGCGTCGTACTGTAAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC 360  
 QY 121 LysValLySLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
 Db 361 AAGGTCAAGTTCAATGGCGTAACCTTCTCCGATGGAACCTGTATGCAAAAGAGACA 420  
 QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLySGlyGlu 160  
 Db 421 ATGGGCTGGGAAGCAAGCACTGAGCGTTGTATCCTCGTATGGCGGTGGAAGAGAG 480  
 QY 161 IleHisLySLeuAlaLeuLySLeuLySAspGlyGlyHisTyrrLeuValGluPheLySerIle 180  
 Db 481 ATTCAATAAGCTCTGAAGCTGAAAGACGGTGTCTATTACCTAGTTGAATTCAAAAGTATT 540

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrrTyrrValAspSerLySLeuAsp 200  
 Db 541 TACATGGCAAAAGAGCTGTGCAGCTACAGGGTACTACTATGTGACTCCAAACTGAT 600  
 QY 201 IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrgLuarGThrGlyArg 220  
 Db 601 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGATAGAAAGAACGAGGAGCG 660  
 QY 221 HisHisLeuPheLeu 225  
 Db 661 CACCATCTGTCTCCTT 675

RESULT 12

US-10-844-064A-1  
 ; Sequence 1, Application US/10844064A  
 ; Publication No. US20050149994A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bevis, Brooke  
 ; TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE S  
 ; FILE REFERENCE: 092234-9006  
 ; CURRENT APPLICATION NUMBER: US/10/844,064A  
 ; PRIORITY FILING DATE: 2004-05-11  
 ; PRIORITY APPLICATION NUMBER: PCT/US02/40539  
 ; PRIORITY FILING DATE: 2002-12-18  
 ; PRIORITY APPLICATION NUMBER: US 60/341,723  
 ; PRIORITY FILING DATE: 2001-12-19  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 678  
 ; TYPE: DNA  
 ; ORGANISM: Discosoma  
 US-10-844-064A-1

Alignment Scores:  
 Pred. No.: 6.83e-143 Length: 678  
 Score: 1214.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-844-064A-1 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
 Db 1 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTTCATGAGGTTTAAGTTTCGATGGAAGCA 60  
 QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
 Db 61 ACGGTCAATGGGCAAGAGTTGAAATAGAACGCGAAGAGAGGGGAGGCCATACGAAGGC 120  
 QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
 Db 121 CACAATACCGTAAGCTTAAGTAACCAAGGGGACCTTGGCATTTGGATATT 180  
 QY 61 LeuSerProGlnPheGlnTyrgLySerLyValTyrrValLyHisProAlaAspIlePro 80  
 Db 181 TTGTCAACCAATTTTCAGTATGGAAGCAAGTATATGTCAAGCACCTTCCGACATACCA 240  
 QY 81 AspTyrrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
 Db 241 GACTATAAAGCTGTCAATTTCTGGAAGATTAAATGGGAAAGGTCATGAACCTTGAA 300  
 QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrr 120  
 Db 301 GACGGTGGCGTGTACTGTAAACCAAGATTCCAGTTTGACAGATGGCTGTTTCATCTAC 360  
 QY 121 LysValLySLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
 Db 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTCCGATGGAACCTGTATGCAAAAGAGACA 420



QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
DB	421	ATGGGCTGGGAAGCCAGCACTGAAGCTTTGTATCCTCGTATGGCGTTGAAGGAGAG	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIle	180
DB	481	ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT	540
QY	181	TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
DB	541	TACATGGCAAGAAGCCCTGTGCAGCTACCAGGGTACTACTATGTTGACTCCAACCTGGAT	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
DB	601	ATAACAAGCCACACAAGAGACTATACATCTGTGACAGATATGAAGAAGACCAGGAGC	660
QY	221	HisHisLeuPheLeu	225
DB	661	CACCATCTGTTCCTT	675

## RESULT 13

```

US-10-931-304-2
/ Sequence 2, Application US/10931304
/ Publication No. US20050196768A1
/ GENERAL INFORMATION:
/ APPLICANT: Taten, Roger
/ APPLICANT: Campbell, Robert
/ APPLICANT: Baird, Geoffrey
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
/ FILE REFERENCE: 39754-0831CP2CP3
/ CURRENT APPLICATION NUMBER: US/10/931,304
/ CURRENT FILING DATE: 2004-08-30
/ PRIOR APPLICATION NUMBER: 10/209,208
/ PRIOR FILING DATE: 2002-07-29
/ PRIOR APPLICATION NUMBER: 10/121,258
/ PRIOR FILING DATE: 2002-04-10
/ PRIOR APPLICATION NUMBER: 09/866,538
/ PRIOR FILING DATE: 2001-05-24
/ PRIOR APPLICATION NUMBER: 09/794,308
/ PRIOR FILING DATE: 2001-02-26
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 678
/ TYPE: DNA
/ ORGANISM: Discosoma sp.
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(678)
/ OTHER INFORMATION: wild-type DsRed
US-10-931-304-2

```

### Alignment Scores:

Alignment Score:	6.83e-143	Length:	678
Pred. No.:	1214.00	Matches:	225
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	9	Gaps:	0
DB:			

US-10-006-922A-12 (1-225) x US-10-931-304-2 (1-678)

QY 1 MecArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20  
|||  
Db 1 ATGAGGCTTCCAAGAATGTTATCAAGAGGTCATGAGTTTAAGGTTCCATGGAAGA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
|||  
Db 61 ACGGTCATGGGCGACGAGTTTGAAATAGAAAGCGAAGAGGGCCATACGAAGGC 120  
QY 41 HisAsnThrValIleLeuIysValThrIleGlyGlyProLeuProPheAlaTyrAspIle 60  
|||

Db	121	CACAATACCGTAAAGCTTAAGGTAA	CCAAAGGGGGACCTTTGCCATTTCCTTGGGATATT	180
QY	61	LeuSerProGlnPheGlnTyrCglYSerIysValTyrValIysHisProAlaSpIlePro		80
Db	181	TTGTCAACCACAATTTCAGTAATGAAGCAAGTATATGTCAAGCACCTTGCCGACATTACCA		240
QY	81	AspTyrIlyslYslLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu		100
Db	241	GACTATAAAAGCTGTCAATTCTCGAAGATTAAATGGAAAGGTCATGAACCTTTGAA		300
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerSerLeuGlnAspGlyCysPheIleTyr		120
Db	301	GACGGTGGCGTCTTACTGTAAACCCAGGATTCCAGTTTGCAAGATGGCTGTTCATCTAC		360
QY	121	LysValIlyspHeileGlyValAsnPheProSerAspGlyProValMetGlnIlyslYsThr		140
Db	361	AAGGTCAAGTTCATTGGCGGTGAACCTTCTCCGATGGAACCTGTTATGCATAAAGAAGACA		420
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu		160
Db	421	ATGGGCTGGGAAGCCAGCACGTGAGCGTTGTATCCCTCGTATGGCGGTGTGAAGAAGAGAG		480
QY	161	IleHisIlyslAlaLeuIlyslLeuYsAspGlyGlyHisIstYrLeuValGluPheIysSerIle		180
Db	481	ATTCTAAGGCTCTGAAGCTGAAGAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT		540
QY	181	TyrMetAlaIlyslYsProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp		200
Db	541	TACATGGCAAGAAGCCCTGTGCAGCTACCAAGGTACTACTATGTTGACTCCAACCTGAT		600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg		220
Db	601	ATAACAAGCCACAACGAAGACTATACATCGTTGAGCAGTATGAAGAACCAGGGAGCGC		660
QY	221	HisHisLeuPheLeu	225	
Db	661	CACCACTCTGTTCCTT	675	

## RESULT 14

```

US-09-999-745-66
; Sequence 66, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-999-745-66

```

### Alignment Scores:

Alignment Score:	9.49e-143	Length:	859
Prod. No.:	1214.00	Matches:	225
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	3	Gaps:	0
DB:			

US-10-006-922A-12 (1-225) x US-09-999-745-66 (1-859)



Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGGCTTCCAGAAGATGTTATCAAGGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 113  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCAAGAGTTTGAATAAGAGCGAAGAGAGGGGAGGCCATACGAAGGC 173  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTGGGATATT 233  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 234 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 293  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAGCTGTCAATTCTCGAAGGATTAAATGGGAAGGGTCATGAACCTTGAA 353  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCTACTGTAAACCAAGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 413  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTCAATTGGCGTGAACTTCTCCGATGGAACCTGTATGCAAAAGAAGACA 473  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCTCGTATGGCGGTGTAAGGAGAG 533  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCTATTACCTAGTGAATTCAAAAGTATT 593  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAACCTGTGCACTACCAAGGTACTACTATGTGACTCCAAACTGGAT 653  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 654 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAGAAGAACCGAGGAGCGC 713  
Qy 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTCTCTT 728  
RESULT 15  
US-09-866-538-11  
; Sequence 11, Application US/09866538  
; Publication No. US20030032088A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/09/866,538  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-09-866-538-11  
Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859

Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0  
US-10-006-922A-12 (1-225) x US-09-866-538-11 (1-859)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGGCTTCCAGAAGATGTTATCAAGAGTTTCATGAGGTTTAAGGTTCCGATGGAAGGA 113  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCAAGAGTTTGAATAAGAGCGAAGAGAGGGGAGGCCATACGAAGGC 173  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 174 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGACCTTGCCATTGCTGGGATATT 233  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 234 TTGTCAACCAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCACCTTGCCGACATACCA 293  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAGCTGTCAATTCTCGAAGGATTAAATGGGAAGGGTCATGAACCTTGAA 353  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCTACTGTAAACCAAGATTCAGTTTGCAAGATGGCTGTTTCATCTAC 413  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 414 AAGGTCAAGTTCAATTGGCGTGAACTTCTCCGATGGAACCTGTATGCAAAAGAAGACA 473  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGTAATCTCGTATGGCGGTGTAAGGAGAG 533  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGCTATTACCTAGTGAATTCAAAAGTATT 593  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAACCTGTGCACTACCAAGGTACTACTATGTGACTCCAAACTGGAT 653  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 654 ATAACAAGCCACAAGAACTATACAATCGTTGAGCAGTATGAAGAAGAACCGAGGAGCGC 713  
Qy 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTCTCTT 728  
RESULT 16  
US-09-794-308-11  
; Sequence 11, Application US/09794308  
; Publication No. US20030170911A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: ZACHARIAS, David  
; APPLICANT: BAIRD, Geoffrey  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530  
; CURRENT APPLICATION NUMBER: US/09/794,308  
; CURRENT FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859



TYPE: DNA  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)..(731)  
US-09-794-308-11

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-794-308-11 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySGluPhenMetArgPhelysValArgMetGluGly 20  
DB 54 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAAGTTCGCAATGGAAGA 113  
QY 21 ThrValAsnGlyHISGluPhenGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCCGAAGAGAGGGGAGGCCATACGAAGC 173  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAGCTTAAGGTAAAGTAACCAAGGGGGACCTTTGCCATTGCTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCGCAATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPhenGlu 100  
DB 294 GACTATAAAAGCTGTCAATTCCTGAAGATTAAATGGGAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
DB 354 GACGGTGGCGTCGTTACTGTAAACCAAGATTCAGTTGCAGAGATGGCTGTTCACTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCCCTCCGATGACCTGTTATGCCAAAAGACACA 473  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 474 ATGGGCTGGGAAGCCAGCACGTGAGCTTGTATCCTCGTATGGCGTGTGMAAGGAGAG 533  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPhenLySLeu 180  
DB 534 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAAGTTCAAAAAGTATT 593  
QY 181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 594 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTATGTGACTCCAAACTGGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATTAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGATGAAGAAGAACCGAGGAGCG 713  
QY 221 HisHisLeuPheLeu 225  
DB 714 CACCATCTGTTCTT 728

RESULT 17  
US-09-865-291-11  
Sequence 11, Application US/09865291  
Publication NO. US20030186229A1  
GENERAL INFORMATION:  
APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
APPLICANT: TSJEN, Roger  
APPLICANT: TING, Alice

APPLICANT: ZHANG, Jin  
TITLE OF INVENTION: EMISSION RADIOMETRIC INDICATORS OF PHOSPHORYLATION  
FILE REFERENCE: REGEN1550  
CURRENT APPLICATION NUMBER: US/09/865,291  
CURRENT FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 859  
TYPE: DNA  
ORGANISM: Discosoma sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)..(731)  
US-09-865-291-11

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-865-291-11 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySGluPhenMetArgPhelysValArgMetGluGly 20  
DB 54 ATGAGGTCTTCCAAAGATGTTATCAAGAGTTCATGAGGTTTAAAGTTCGCAATGGAAGA 113  
QY 21 ThrValAsnGlyHISGluPhenGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 114 ACGGTCAATGGGCACGAGTTTGAATAGAAAGCCGAAGAGAGGGGAGGCCATACGAAGC 173  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAGCTTAAGGTAAAGTAACCAAGGGGGACCTTTGCCATTGCTGGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCTGCGCAATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPhenGlu 100  
DB 294 GACTATAAAAGCTGTCAATTCCTGAAGATTAAATGGGAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCySPheIleTyr 120  
DB 354 GACGGTGGCGTCGTTACTGTAAACCAAGATTCAGTTTGCAAGATGGCTGTTCACTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCCCTCCGATGACCTGTTATGCCAAAAGACACA 473  
QY 141 MetGlyTTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 474 ATGGGCTGGGAAGCCAGCACGTGAGCTTGTATCCTCGTATGGCGTGTGAAAGGAGAG 533  
QY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPhenLySLeu 180  
DB 534 ATTCAATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAAGTTCAAAAAGTATT 593  
QY 181 TyrMetAlaLySLeuSProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 594 TACATGGCAAGAAGCCTGTGCAGCTACCAAGGTACTATGTGACTCCAAACTGGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATTAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGATGAAGAAGAACCGAGGAGCG 713  
QY 221 HisHisLeuPheLeu 225  
DB 714 CACCATCTGTTCTT 728



RESULT 18  
US-10-433-640-12  
; Sequence 12, Application US/10433640  
; Publication No. US20040115792A1  
; GENERAL INFORMATION:  
; APPLICANT: Lichtenberg-Frate, Hella  
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF  
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS  
; FILE REFERENCE: 1487/3  
; CURRENT APPLICATION NUMBER: US/10/433,640  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/14610  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: DE 10061872.3  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 12  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-10-433-640-12

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-433-640-12 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 54 ATGAGGCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
DB 114 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAGCTTAAAGTAACCAAGGGGGACCTTGCCATTGCTGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTCAGTATGGAAGCAAGATATGTCAAGCAACCCGCGCATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
DB 294 GACTATAAAAAGCTGTCATTTCTGAAAGATTAAATGGAAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGTGCGCTGTTACTGTAAACCCAGATTCAGTTTGAGGATGGCTGTTTCATCTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThr 140  
DB 414 AAGGTCAAGTTCAATGGCGTGAACCTTCCGATGAGCCTGTTATGCAAAAGAAGACA 473  
QY 141 MetGlyTrrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 474 ATGGCGTGGGAAGCAGCACTGAGCGTTGTATCCTCGTATGGCGGTGAAAGAGAG 533  
QY 161 IleHisLySLeuLySLeuLySLeuLySLeuLySLeuLySLeuLySLeuLySLeu 180  
DB 534 ATTCATTAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593

QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrValAspSerLySLeuAsp 200  
DB 594 TACATGGCAAGAAGCCTGTGACGTACCAAGGTAATACTACTATGTTGAATCCAAACTGAT 653  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 654 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCG 713  
QY 221 HisIleLeuPheLeu 225  
DB 714 CACCATCTGTTCTT 728

RESULT 19  
US-10-885-988-11  
; Sequence 11, Application US/10885988  
; Publication No. US20040259165A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: TSJEN, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS  
; FILE REFERENCE: REGEN1530-2  
; CURRENT APPLICATION NUMBER: US/10/885,988  
; CURRENT FILING DATE: 2004-07-06  
; PRIOR APPLICATION NUMBER: US/09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-10-885-988-11

Alignment Scores:  
Pred. No.: 9.49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-885-988-11 (1-859)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 54 ATGAGGCTTCCAAAGATGTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGGA 113  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40  
DB 114 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGCGAAGAGAGGGGAGGCCATACGAAGGC 173  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 174 CACAATACCGTAAGCTTAAAGTAACCAAGGGGGACCTTGCCATTGCTGGATATT 233  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHisProAlaAspIlePro 80  
DB 234 TTGTCAACCAATTCAGTATGGAAGCAAGTATATGTCAAGCAACCCGCGCATACCA 293  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrgLuarGValMetAsnPheGlu 100  
DB 294 GACTATAAAAAGCTGTCATTTCTGAAAGATTAAATGGGAAGGTCATGAACCTTGAA 353  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 354 GACGGTGCGCTGTTACTGTAAACCCAGATTCAGTTTGAGGATGGCTGTTTCATCTAC 413  
QY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySlySThr 140



Db 414 AAGTCAGTTCAATGGCGTGAACTTCTTCGATGGACCTGTTATGCAAGAAGACA 473  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCAGCACTGAGCGTTTGATCCTCGTATGGCGGTGTGAAGAAGAG 533  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAAGACCTGTGCAAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 653  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlyArg 220  
Db 654 ATAACAAGCCACACAGCAAGACTATACAATCGTTGAGCAGATATGAAGAAGCCAGGAGCGC 713  
Qy 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCCCTT 728

RESULT 20  
US-10-857-622-11  
; Sequence 11, Application US/10857622  
; Publication No. US20050026234A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: VIOLIN, Jonathan  
; APPLICANT: NEWTON, Alexandra  
; APPLICANT: TSIEH, Roger  
; APPLICANT: ZHANG, Jin  
; TITLE OF INVENTION: Emission Ratiometric Indicators of Phosphorylation By C-Kinase  
; FILE REFERENCE: 39754-0891 CPC1CP2  
; CURRENT APPLICATION NUMBER: US/10/857,622  
; CURRENT FILING DATE: 2004-05-28  
; PRIOR APPLICATION NUMBER: US 09/865,291  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: US 09/396,003  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: US 08/792,553  
; PRIOR FILING DATE: 1997-01-31  
; PRIOR APPLICATION NUMBER: US 594,575  
; PRIOR FILING DATE: 1996-01-31  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (54)..(731)  
US-10-857-622-11

Alignment Scores:  
Pred. No.: 9,49e-143 Length: 859  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-857-622-11 (1-859)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 54 ATGAGTCTTCCAAGATGTATCAAGAGTTCATGAGTTTAAGTTTCGCATGAAGGA 113  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 114 ACGGTCAATGGGCAAGAGTTTGAATATGAAGCGGAAGAGAGGAGGCCATACGAAGGC 173

Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProDheAlaTrpAspIle 60  
Db 174 CACAATACCGTAAGCTTAAGTAACCAAGGGGAGCCTTGGCAATTGCTGGATATT 233  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 234 TTGTACCAACAATTTCAGTATGGAAGCAAGTATATGTCAAGCACCTGCCGACATACCA 293  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 294 GACTATAAAAGCTGTCAATTCTTGAAAGATTAAATGGGAAAGGTCATGAACCTTGAA 353  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 354 GACGGTGGCGTCTTAAGTGAACCAAGATTCCAGTTGACAGATGGCTGTTCACTAC 413  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 414 AAGTCAAGTTCATTGGCGTGAACTTCTTCGATGGAACCTGTTATGCAAAAGAAGACA 473  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 474 ATGGGCTGGGAAGCCAGCACTGAGCGTTTGATCCTCGTATGGCGGTGTGAAGAAGAG 533  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 534 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 593  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 594 TACATGGCAAGAAGACCTGTGACAGCTACCAAGGTACTACTATGTGACTCCAAACTGGAT 653  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlyArg 220  
Db 654 ATAACAAGCCACACAGCAAGACTATACAATCGTTGAGCAGATATGAAGAAGCCAGGAGCGC 713  
Qy 221 HisHisLeuPheLeu 225  
Db 714 CACCATCTGTTCCCTT 728

RESULT 21  
US-09-797-496B-3  
; Sequence 3, Application US/09797496B  
; Publication No. US20030049597A1  
; GENERAL INFORMATION:  
; APPLICANT: Simon, Sanford M.  
; APPLICANT: Chen, Yu  
; TITLE OF INVENTION: Chimeric Fluorescent Enzymes and Uses Thereof  
; FILE REFERENCE: 600-1-267  
; CURRENT APPLICATION NUMBER: US/09/797,496B  
; CURRENT FILING DATE: 2002-05-24  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 3311  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Discosoma red fluorescent protein modified as described in specif  
; OTHER INFORMATION: ication.  
US-09-797-496B-3

Alignment Scores:  
Pred. No.: 6,22e-142 Length: 3311  
Score: 1214.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-006-922A-12 (1-225) x US-09-797-496B-3 (1-3311)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20



|||||  
Db 289 ATGAGGCTTCCAGAATGTTATCAAGAGATTCAAGGTTTAAGGTTCCATGGAAGA 348  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 349 ACGGTCAATGGGCAAGAGTTTGAATAGAAAGGCGAAGAGAGGGGAGCCATACGAAGGC 408  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 409 CACAATACCGTAAGCTTAAGGTAACCAAGGGGGGACCTTGGCCATTGGTTGGATATT 468  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 469 TTGTCACCACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCA 528  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 529 GACTATAAAAGCTGTCTATTCTCGAAGGATTTAATGGGAAAGGGTCTGAACCTTGAA 588  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 589 GACGGTGGCGTCTTACTGTAACCCAGGATTCCAGTTTGCAAGATGGCTGTTTCATCTAC 648  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 649 AAGGTCAAGTTCAATTGGCGTGAACTTTCCTCCGATGGACCTGTATGCAAAAGAAACA 708  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 709 ATGGCGTGGGAAGCCAGCACTGAGCGTTTGTAATCCTCGTATGGCGGTGTGAAGAAGAG 768  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 769 ATTCATAAGGCTCTGAAGCTGAAGAAGCGGTGTCATTACCTAGTTGAATTCAAAAGTATT 828  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 829 TACATGGCAAGAACGCTGTGCAGCTACCGAGGTACTACTATGTGACTCCAAACTGGAT 888  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 889 ATAACAAGCCACAACGAAGACTATACAATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC 948  
Qy 221 HisHisLeuPheLeu 225  
Db 949 CACCATCTGTTCTT 963  
RESULT 22  
US-10-006-922-35  
; Sequence 35, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/457,556  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/444,338

; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Discosoma species  
US-10-006-922-35  
Alignment Scores:  
Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 5 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-006-922-35 (1-681)  
Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 4 GTGCGCTCTCCAAAGACGTCTCAAGAGTTTCATGCGCTTCAAGTGCCGATGAGAGGC 63  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCGGCCCTACGAGGC 123  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 124 CACAACACCGTGAAGCTGAAGGTGAACCAAGGGCGGCCCTGCCCTTCGCTGGACATC 183  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 184 CTGTCCCCCAGTTCCAGTACGGCTCCAAGGTGACGTGAAGCAACCCCGACATCCCC 243  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
Db 244 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGAATTCGAG 303  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 304 GACGGCGCGTGTGACCGCTGACCACAAGACTCTCCCTGCAAGACGGCTGCTTATCTAC 363  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 364 AAGTGAAGTTCAATCGCGGTGAATTCCTCCGACGGCCCGTAATGCAAGAAGACC 423  
Qy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 424 ATGGCGTGGGAGGCTCTCACCGAGGCTGTACCCCGCGAGCGGTGCTGAAGGGCGAG 483  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGGAGTTCAAGTCCATC 543  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 544 TACATGGCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTGAGACTCCAAAGCTGGAC 603  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 604 ATCACCTCCACAACGAGGACTACACATCGTGAGCAGTACGAGCGCACCGAGGGCGGC 663  
Qy 221 HisHisLeuPheLeu 225  
Db 664 CACCACCTGTTCTTG 678  
RESULT 23  
US-10-121-258-3  
; Sequence 3, Application US/10121258  
; Publication No. US20030059835A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger



APPLICANT: Campbell, Robert  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/10/121,258  
CURRENT FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 681  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
OTHER INFORMATION: codon usage  
US-10-121-258-3

Alignment Scores:  
Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-3 (1-681)

OY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 4 GTGCGCTCCTCCAGAACGTATCAAGAGTTCATGCGCTTCAAGTGCGCATGAGGCG 63  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGGAGGCGGCGCCCTACGAGGCG 123  
OY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTTCCGCTGGGACATC 183  
OY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGATGAAAGCAACCCGCGACATCCCC 243  
OY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG 303  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 304 GACGGCGGCGGTGTGACCGGTGACCCAGGACTCTCTGCAAGCGGCTGCTTCTTAC 363  
OY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 364 AAGGTGAAGTTCAATCGGCGTGAATTCCCTCCGACGCGCGCTGATGCAAGAAAGACC 423  
OY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGGAG 483  
OY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuThr 180  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGAGCGGCGGCACTACCTGGTGAAGTTCACATC 543  
OY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
DB 544 TACATGGCCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGAAGTCCAAAGCTGAG 603  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 604 ATCACCCTCCACAACGAGACTACCAATCGTGAGCAGTACGAGCGCACCGAGGGCCGC 663  
OY 221 HisIleuPheLeu 225  
DB 664 CACCACCTGTTCTCG 678

RESULT 24  
US-10-121-258-23

; Sequence 23, Application US/10121258  
; Publication No. US20030059835A1

; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger

; APPLICANT: Campbell, Robert

; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT

; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

; FILE REFERENCE: UC083.1CP2CP1

; CURRENT APPLICATION NUMBER: US/10/121,258

; CURRENT FILING DATE: 2002-04-10

; PRIOR APPLICATION NUMBER: 09/794,308

; PRIOR FILING DATE: 2001-02-26

; PRIOR APPLICATION NUMBER: 09/866,538

; PRIOR FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 78

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 23

; LENGTH: 681

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian

; OTHER INFORMATION: codon usage  
US-10-121-258-23

Alignment Scores:  
Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-23 (1-681)

OY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
DB 4 GTGCGCTCCTCCAGAACGTATCAAGAGTTCATGCGCTTCAAGTGCGCATGAGGCG 63  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGGAGGCGGCGCCCTACGAGGCG 123  
OY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCGCCCTGCTTCCGCTGGGACATC 183  
OY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGATGAAAGCAACCCGCGACATCCCC 243  
OY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAGAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG 303  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 304 GACGGCGGCGGTGTGACCGGTGACCCAGGACTCTCTGCAAGAGCGGCTGCTTCTTAC 363  
OY 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
DB 364 AAGGTGAAGTTCAATCGGCGTGAATTCCCTCCGACGCGCGCTGATGCAAGAAAGACC 423  
OY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160



Db 424 ATGGGCTGGAGGCGCTCCACCGAGCGCCTGTACCCCGCGACGGCGCTGTGAAGGCGAG 483  
QY 161 ILeHsLysAlaLeuLyLeuLySAAspGlyGlyHISLysLeuValGluPheLysSerile 180  
Db 484 ATCCACAAGGCGCTGAAGCTGAAGACGGCGGCCACTACCTGCTGAGTTCAGATCCATC 543  
QY 181 TyrMetAlaLysLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 544 TACATGGCCCAAGAACCCGTGACAGCTGCCGCTACTACTAGTGAAGCTCCAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluYArg 220  
Db 604 ATCACCTCCCAACAAGGAGCTACACCATCTGTGAGCAAGTACGAGCGCACCGAGGGCCGC 663  
QY 221 HIsHsLeuPheLeu 225  
Db 664 CACCACCTGTTCTCTG 678

RESULT 25

US-10-311-030-8

; Sequence 8, Application US/10311030  
; Publication No. US20040171107A1  
; GENERAL INFORMATION:  
; APPLICANT: Nelson, David  
; APPLICANT: Zamiatra, Elize  
; APPLICANT: Tsien, Roger  
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS  
; FILE REFERENCE: 15916-032US1  
; CURRENT APPLICATION NUMBER: US/10/311,030  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: PCT/US01/04625  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: US 60/184,732  
; PRIOR FILING DATE: 2000-02-23  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(678)  
US-10-311-030-8

Alignment Scores:

Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-8 (1-681)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 4 GTGAGGAGCAGCAAGAACGTGATCAAGAGTTTCATGAGTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 124 CACAACACCGTGAAGCTTAAGGTGACCAAGGGCGGCCCTGCTCGCTGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 184 CTGAGCCCCCACTTCAGTACGCGCAGCAAGGTGTACGTGAAGCACCCCGCGACATCCCC 243

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 244 GACTACAGAAGAGCTGAGCTTCCCCAGGGCTTCAAGTGGAGAGGGTGATGAACCTCGAG 303  
QY 101 AspGlyGlyValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 304 GACGGCGCGGTGTGACCGGTGACCAAGACAGACAGCCTGCAAGACGGCTGCTCATCTAC 363  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 364 AAGGTGAAGTTCATCGGCGTGAAGTTCACCAAGCAGCGGCCCGGTGATGCAAGAAGACC 423  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 483  
QY 161 ILeHsLysAlaLeuLyLeuLySAAspGlyGlyHISLysLeuValGluPheLysSerile 180  
Db 484 ATCCACAAGGCGCTGAAGCTGAAGACGGCGGCACCTACTGTGTGAGTTCAGATCCATC 543  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 544 TACATGGCCCAAGAACCCGTGACAGCTGCCGCTACTACTAGTGAAGTCCAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluYArg 220  
Db 604 ATCACCAAGCCACAAGAGGAGCTACACCATCTGTGAGAGCTACGAGAAGACCGAGGGCAGG 663  
QY 221 HIsHsLeuPheLeu 225  
Db 664 CACCACCTGTTCTCTG 678

RESULT 26

US-10-931-304-3

; Sequence 3, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-10-931-304-3

Alignment Scores:

Pred. No.: 2.2e-142 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-3 (1-681)



```
QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      4 GTCCGCTCCTCCAGAGACGTCAACAGAGAGTCAAGCTTCAAGTGCAGTGAAGGCG 63
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      64 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCCAGGGCCGCCCTTACGAGGGC 123
QY      41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheIleTrpAspIle 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      124 CACAACACCGTGAAGGTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTTGGGACATC 183
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProIleAspIlePro 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      184 CTGTCCCCCAGTTCCAGTACGGCTCCAGAGTGTACGTAAGCACCCCGCCGACATCCCC 243
QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySerTrpGluArgValMetAsnPheGlu 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      244 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTCGAG 303
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      304 GACGGCGGCGTGTGACCGGTGACCCAGACTCTCTCCCTGCAAGGACGGCTGTCACTAC 363
QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      364 AAGGTGAAGTTCAATCGGCTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAGAAC 423
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      424 ATGGGCTGGAGGCTCCAGCGGAGCGGCTGTACCCCGCGACGGCGTGTGAAGGGGAG 483
QY      161 IleHisLySAlaLeuLyLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      484 ATCCACAAGGCCCTGAAGCTGAAGAGCGGGCCCACTACCTGGAGTTCAAGTCCATC 543
QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTGAGTCCAAAGCTGAC 603
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      604 ATCACTCCCAACAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCG 663
QY      221 HisHisLeuPheLeu 225
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      664 CACCACTGTTCCTG 678

RESULT 27
US-10-931-304-23
; Sequence 23, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Teslen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
```

```
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian
; OTHER INFORMATION: codon usage
US-10-931-304-23

Alignment Scores:
Pred. No.:      2,2e-142      Length:      681
Score:          1210.00      Matches:      224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches: 0
Query Match:    99.67%      Indels:      0
DB:              9      Gaps:      0

US-10-006-922A-12 (1-225) x US-10-931-304-23 (1-681)

QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      4 GTCCGCTCCTCCAGAGACGTCAACAGAGAGTCAAGCTTCAAGTGCAGTGAAGGCG 63
QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      64 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCCAGGGCCGCCCTTACGAGGGC 123
QY      41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheIleTrpAspIle 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      124 CACAACACCGTGAAGGTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTTGGGACATC 183
QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProIleAspIlePro 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      184 CTGTCCCCCAGTTCCAGTACGGCTCCAGAGTGTACGTAAGACACCCCGCCGACATCCCC 243
QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySerTrpGluArgValMetAsnPheGlu 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      244 GACTACAAGAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 303
QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      304 GACGGCGGCGTGTGACCGGTGACCCAGACTCTCTCCCTGCAAGAGCGGCTGTCACTAC 363
QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      364 AAGGTGAAGTTCAATCGGCTGAACCTTCCCTCCGACGGCCCCGTAATGCAAGAGAAC 423
QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      424 ATGGGCTGGAGGCTCCAGCGGAGCGGCTGTACCCCGCGAGCGGCTGTGAAGGGCGAG 483
QY      161 IleHisLySAlaLeuLyLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      484 ATCCACAAGGCCCTGAAGCTGAAGAGCGGGCCCACTACCTGGAGTTCAAGTCCATC 543
QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTGAGTCCAAAGCTGAGC 603
QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      604 ATCACTCCCAACAGAGACTACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCG 663
QY      221 HisHisLeuPheLeu 225
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      664 CACCACTGTTCCTG 678

RESULT 28
US-10-311-030-11
; Sequence 11, Application US/10311030
; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zamiatara, Elize
```



; APPLICANT: Tsien, Roger
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
; FILE REFERENCE: 15916-032US1
; CURRENT APPLICATION NUMBER: US/10/311,030
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US01/04625
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/184,732
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)...(696)
; US-10-311-030-11

Alignment Scores:
Pred. No.: 2.34e-142 Length: 713
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-11 (1-713)

QY 1 MetArgSerSerLySAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 22 GTGAGGAGCAGCAAGAACTGATCAAGAGTTTCATGAGTTCAAGGTGCGCATGAGGGC 81
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB 82 ACCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGAGGGCGAGGCCCTTACGAGGGC 141
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
DB 142 CACAACACCGTGAAGCTTAAGTGACCAAGGGCGGCCCCCTGCGCTGGGACATC 201
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
DB 202 CTGAGCCCCCAGTTCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCCGCACATCCCC 261
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
DB 262 GACTACAAAGAGTGAAGCTTCCCCGAGGGCTTCAAGTGGAGAGGGTGAATGAACCTTCGAG 321
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 322 GACGGCGGCGTGTGACCGTGACCCAGCAGACGCTGCAGAGCGGCTGCTTCATCTAC 381
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB 382 AAGGTGAAGTCAATCGCGGTGAACCTCCCAAGCGGCGGCGGCTGATGCAAGAAGACC 441
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB 442 ATGGGCTGGGAGGCTCCACCGAGCGGCTGTACCCCGGACGGCGGTGTGAAGGGCGAG 501
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 502 ATCCACAAGGCCCTGAAGCTGAAGGAGCGGGGCCACTACCTGTGTGAGTTCAAGTCCATC 561
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB 562 TACATGGCCAAAGAACCCCGTGCAAGCTGCCGCGCTACTACTACGTGAGCTCCAAAGCTGAC 621
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 622 ATCACCAGCCACAACGAGAGTACCATCTGTGAGCAGTACGAGAGACCGAGGGCAGG 681
QY 221 HisHisLeuPheLeu 225
DB 682 CACCACCTGTTCCTG 696

RESULT 29
US-10-311-030-12/c
; Sequence 12, Application US/10311030
; Publication No. US20040171107A1
; GENERAL INFORMATION:
; APPLICANT: Nelson, David
; APPLICANT: Zamiatra, Elize
; APPLICANT: Tsien, Roger
; TITLE OF INVENTION: MODIFIED FLUORESCENT PROTEINS
; FILE REFERENCE: 15916-032US1
; CURRENT APPLICATION NUMBER: US/10/311,030
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US01/04625
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: US 60/184,732
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; US-10-311-030-12

Alignment Scores:
Pred. No.: 2.34e-142 Length: 713
Score: 1210.00 Matches: 224
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 0
Query Match: 99.67% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-311-030-12 (1-713)

QY 1 MetArgSerSerLySAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB 692 GTGAGGAGCAGCAAGAACTGATCAAGAGTTTCATGAGTTCAAGGTGCGCATGAGGGC 633
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB 632 ACCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGAGGGCGAGGCCCTTACGAGGGC 573
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
DB 572 CACAACACCGTGAAGCTTAAGTGACCAAGGGCGGCCCCCTGCGCTGGGACATC 513
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
DB 512 CTGAGCCCCCAGTTCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCCGCACATCCCC 453
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
DB 452 GACTACAAAGAGCTGAAGCTTCCCCGAGGGCTTCAAGTGGAGAGGGTGAATGAACCTTCGAG 393
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB 392 GACGGCGGCGTGTGACCGTGACCCAGCAGACGACCTGCAGAGCGGCTTCATCTAC 333
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB 332 AAGGTGAAGTTCATCGCGGTGAACCTCCCAAGCGGCGGCGGCTGATGCAAGAAGACC 273
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160



Db 272 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCCGCGAGCGGCTGTGAAGGGCGAG 213  
Qy 161 ILeHsIySAlaLeuLySLeuLySAspGlyGlyHISlyrLeuValGluPheLySserIle 180  
Db 212 ATCCACAGGCCCTGAAGCTGAAGGACGGCGCCACTACCTGTGAGTTCAAGTCCATC 153  
Qy 181 TyrMeAlaLySLeuProValGluLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 152 TACATGGCCCAAGAGCCCGTGCACTGCGGCTACTACTAGTGAAGTCCAAAGCTGAGC 93  
Qy 201 IleThrSerHISaNGluAspTyrThrIleValGluGluIntYrGluArgThrGluGlyArg 220  
Db 92 ATCACCAGCCCAAGAGGACTACACCATCTGTGAGCAGTACGAGAGACCGAGGGCAGG 33  
Qy 221 HIsHISLeuPheLeu 225  
Db 32 CACCACCTGTTCTCTG 18

RESULT 30

US-10-152-296-1  
; Sequence 1, Application US/10152296  
; Publication No. US20030077730A1  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/152,296  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 723  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(723)  
; OTHER INFORMATION: DsRED  
; US-10-152-296-1

Alignment Scores:  
Pred. No.: 2.39e-142 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
Gaps: 0  
DB:

US-10-006-922A-12 (1-225) x US-10-152-296-1 (1-723)

Qy 1 MetArgSerSerLySAsnValIleLySGLuPheMetArgPheLySValArgMetGluGly 20  
Db 4 GTGCGCTCCTCCAAAGACGTCAATCAAGAGTTCAATCGCTTCAAGGTGCGCATGAGGGC 63  
Qy 21 ThrValaNGlyHISGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 64 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC 123  
Qy 41 HIsaSnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
Db 124 CACAACAACCGTGAAAGTGGAAGCAAGGGCGGCGCCCTTCGCTGGGACATC 183  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80  
Db 184 CTGTCCCGCCAGTTCAGTACGGCTCCAAGGTGTATGTAAGCAACCGCGGACATCCCC 243

Qy 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 244 GACTACAGAAGCTGTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 303  
Qy 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
Db 304 GACGGCGCGTGCTGACCGCTGACCCAGGACTCCTCCCTGCAGAGACGGCTTCATCTAC 363  
Qy 121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 364 AAGGTGAATTCAATCGGCGTGAATTCTCCCTCCGACGGCCCCGTATATCAGAAGAAGACC 423  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160  
Db 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 483  
Qy 161 ILeHsIySAlaLeuLySLeuLySAspGlyGlyHISlyrLeuValGluPheLySserIle 180  
Db 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGCCACTACCTGTGAGTTCAAAGATATC 543  
Qy 181 TyrMeAlaLySLeuProValGluLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 544 TACATGGCCCAAGAGCCCGTGCACTGCGGCTACTACTAGTGAAGTCCAAAGCTGAGC 603  
Qy 201 IleThrSerHISaNGluAspTyrThrIleValGluGluIntYrGluArgThrGluGlyArg 220  
Db 604 ATCACCCTCCCAAGAGGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 663  
Qy 221 HIsHISLeuPheLeu 225  
Db 664 CACCACCTGTTCTCTG 678

RESULT 31

US-10-739-656-1  
; Sequence 1, Application US/10739656  
; Publication No. US20040126850A1  
; GENERAL INFORMATION:  
; APPLICANT: Peelle, Beau  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Directed Evolution of Protein in Mammalian Cells  
; FILE REFERENCE: 021044-000110US  
; CURRENT APPLICATION NUMBER: US/10/739,656  
; PRIOR FILING DATE: 2003-12-17  
; PRIOR APPLICATION NUMBER: US/10/152,296  
; PRIOR FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/291,871  
; PRIOR FILING DATE: 2001-05-18  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 723  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:mammalian  
; OTHER INFORMATION: codon-optimized variant (DsRED) of Discosoma sp.  
; OTHER INFORMATION: "red" red fluorescent protein (RFP)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(723)  
; OTHER INFORMATION: DsRED  
; US-10-739-656-1

Alignment Scores:  
Pred. No.: 2.39e-142 Length: 723  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
Gaps: 0  
DB:

US-10-006-922A-12 (1-225) x US-10-739-656-1 (1-723)



```

Qy      1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20
      :::|||||
Db      4 GTGCGCTCTCCAGAAGCGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63

Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||||
Db      64 ACCGTGAACGGCCACAGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGGC 123

Qy      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
      |||||
Db      124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTCGGACATC 183

Qy      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80
      |||||
Db      184 CTGTCCCCCAAGTTCAGTAGCGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 243

Qy      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||||
Db      244 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 303

Qy      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
      |||||
Db      304 GACGGCGGGGTGTGACCGTGACCCAGACTCTCCCTGCAGAGCGGCTTTCATCTAC 363

Qy      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      |||||
Db      364 AAGGTGAAGTTCATCGGCGTGAACTTCCCTCCGACGGCCCCGTATGCAGAAGAAGACC 423

Qy      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      |||||
Db      424 ATGGGCTGGAGGCTTCCACCGAGCGGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 483

Qy      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
      |||||
Db      484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCCACTACTGTGTGAGTTCAAGATATC 543

Qy      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      |||||
Db      544 TACATGGCCAGAAGCCCGTGACAGCTGCCGGCTACTACTACGTGACTCCAAAGCTGAGC 603

Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      |||||
Db      604 ATCACCCTCCACAAGAGACTACACCATCGTGAAGCAGTAGAGCGCACCGAGGGCGGC 663

Qy      221 HisHisLeuPheLeu 225
      |||||
Db      664 CACCACCTGTTCCTG 678

RESULT 32
US-10-214-932-51
; Sequence 51, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: AP802/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1638)
; OTHER INFORMATION: Gene for RFP:PS(N1A protease):AtOEP7:GFP
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (1)..(1635)
; OTHER INFORMATION: RFP:PS (N1A protease):AtOEP7:GFP hybrid protein
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: CDS: RFP(1..678); CDS: N1A protease proteolytic site(700..711); C
; OTHER INFORMATION: DS: AtOEP7(712..900); CDS: GFP(922..1635)
US-10-214-932-51

Alignment Scores:
Pred. No.:      7.46e-142      Length:      1638
Score:          1210.00      Matches:      224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches: 0
Query Match:     99.67%      Indels: 0
DB:              5      Gaps: 0

US-10-006-922A-12 (1-225) x US-10-214-932-51 (1-1638)

Qy      1 MetArgSerSerLySAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20
      :::|||||
Db      4 GTGCGCTCTCCAGAAGCGTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63

Qy      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||||
Db      64 ACCGTGAACGGCCACAGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTACGAGGGC 123

Qy      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
      |||||
Db      124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTTCGGACATC 183

Qy      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLySHisProAlaAspIlePro 80
      |||||
Db      184 CTGTCCCCCAAGTTCAGTAGCGGCTCCAAGGTGTACGTGAAGCACCCCGCCGACATCCCC 243

Qy      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||||
Db      244 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 303

Qy      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
      |||||
Db      304 GACGGCGGGGTGTGACCGTGACCCAGACTCTCCCTGCAGAGCGGCTTTCATCTAC 363

Qy      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140
      |||||
Db      364 AAGGTGAAGTTCATCGGCGTGAACTTCCCTCCGACGGCCCCGTATGCAGAAGAAGACC 423

Qy      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      |||||
Db      424 ATGGGCTGGAGGCTTCCACCGAGCGGCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 483

Qy      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180
      |||||
Db      484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 543

Qy      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      |||||
Db      544 TACATGGCCAGAAGCCCGTGACAGCTGCCGGCTACTACTACGTGACTCCAAGCTGAGC 603

Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      |||||
Db      604 ATCACCCTCCACAAGAGACTACACCATCGTGAAGCAGTAGAGCGCACCGAGGGCGGC 663

Qy      221 HisHisLeuPheLeu 225
      |||||
Db      664 CACCACCTGTTCCTG 678

RESULT 33
US-10-214-932-75
; Sequence 75, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan

```



```

/ APPLICANT: KIM, Dae Heon
/ APPLICANT: LEE, Yong Jik
/ TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
/ FILE REFERENCE: APB02/US
/ CURRENT APPLICATION NUMBER: US/10/214,932
/ CURRENT FILING DATE: 2002-08-08
/ NUMBER OF SEQ ID NOS: 133
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 75
/ LENGTH: 1647
/ TYPE: DNA
/ ORGANISM: Artificial
/ FEATURE:
/ OTHER INFORMATION: Synthetic Sequence
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (1)..(1647)
/ OTHER INFORMATION: Gene for RFP:PS (HIV-1 protease):AtOEP7:GFP hybrid protein
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(1644)
/ OTHER INFORMATION: RFP:PS (HIV-1 protease):AtOEP7:GFP hybrid protein
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: CDS: RFP(1..678) ; CDS: HIV-1 protease proteolytic site(700..720) ;
/ OTHER INFORMATION: CDS: AtOEP7(721..909) ; CDS: GFP(931..1644)
US-10-214-932-75

```

Alignment Scores:

Pred. No.:	7.52e-142	Length:	1647
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	5	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-214-932-75 (1-1647)

```

QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly  20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      4 GTGGCGTCTCTCCAAGAACGTCAAGAGTTTCATGCGCTTCAAGTGCGCATGAGGCG  63

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly  40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      64 ACCGTGAACGGCCACGAGTTCGAGTCGAGGGCGGAGGGCGCCGCCCTACGAGGCG  123

QY      41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle  60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      124 CACAACAACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTGGAGCATC  183

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro  80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      184 CTGTCCCCCAGTTCAGTACGGCTCCAAAGGTGATGAGCAACCCGCCGACATCCCC  243

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu  100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      244 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCAG  303

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr  120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      304 GACGGCGCGTGGTGAACCGTGAACCGAGACTCTCCCTGCAGAGCGGCTGCTTCATCTAC  363

QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr  140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      364 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCCGTAATGCAAGAAGACC  423

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu  160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      424 ATGGGCTGGAGGGCTTCACCGAGCGCTGTACCCCCCGGACGGCGCTGTAAGGGCGAG  483

QY      161 IleHisLySValLeuLySLeuLySLeuLySLeuLySLeuLySLeuLySLeuLySLeu  180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      484 ATCCACAAGGCCCTGAAGCTGAAGGAGCGCGGCCAATACTGTGGAGTTCAAGTCCATC  543

```

```

QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp  200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      544 TACATGGCCAGAAGCCCGTGACGCTGCCGCGCTACTACTACGTGACTCCAAAGCTGAC  603

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg  220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      604 ATCACCTCCCAACAAGAGACTACACCATGTGAGCAGTACGAGCGCACCGAGGGCCGC  663

QY      221 HisHisLeuPheLeu  225
      ::::::::::::::::::::
DB      664 CACCACCTGTTCCTG  678

```

RESULT 34

```

US-10-161-403-29
/ Sequence 29, Application US/10161403
/ Publication No. US20030119104A1
/ GENERAL INFORMATION:
/ APPLICANT: Perkins, Edward
/ APPLICANT: Perez, Carl
/ APPLICANT: Lindenbaum, Michael
/ APPLICANT: Greene, Amy
/ APPLICANT: Leung, Josephine
/ APPLICANT: Fleming, Elena
/ APPLICANT: Stewart, Sandra
/ APPLICANT: Shellard, Joan
/ TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
/ FILE REFERENCE: 24601-420
/ CURRENT APPLICATION NUMBER: US/10/161,403
/ CURRENT FILING DATE: 2002-05-30
/ PRIOR APPLICATION NUMBER: 60/294,758
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: 60/366,891
/ PRIOR FILING DATE: 2002-03-21
/ NUMBER OF SEQ ID NOS: 129
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29
/ LENGTH: 4692
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: pDsred1-N1 plasmid from Clontech
US-10-161-403-29

```

Alignment Scores:

Pred. No.:	3.23e-141	Length:	4692
Score:	1210.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	0
Query Match:	99.67%	Indels:	0
DB:	6	Gaps:	0

US-10-006-922a-12 (1-225) x US-10-161-403-29 (1-4692)

```

QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly  20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      682 GTGGCGTCTCTCCAAGAACGTCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGCG  741

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly  40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      742 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGGAGGGCGGCCCTTGCCTGGACATC  801

QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle  60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      802 CACAACAACCGTGAAGCTGAAGTGAACCAAGGGCGGCCCTTGCCTGGAGCATC  861

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro  80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      862 CTGTCCCCCAGTTCAGTACGGCTTCAAGGTGATGTAAGCAACCCCGGACATCCCC  921

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu  100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      922 GACTACAAGAAGCTGTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG  981

```



Qy 101 AspglyglyValValThrValThrGlnAspSerSerLeuGlnAspglyCysPheIleTyr 120  
Db 982 GACGGCGCGGTGGTGACCGGTGACCCAGGACTCTCTCGACAGACGGCTGCTCATCTAC 1041  
Qy 121 LysValLysPheIleGlyValAsnProSerAspglyProValMetGlnLysLysThr 140  
Db 1042 AAGGTGAAGTTTCATCGCGGTGAACCTCCCTCCGACGGCCCCGTAAATGCAAGAAAGACC 1101  
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspglyValLeuLysGlyGlu 160  
Db 1102 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGCTGTGAAGGGCGAG 1161  
Qy 161 IleHisLysAlaLeuLysLeuLysAspglyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1162 ATCCACAAGGCCCTGAAGCTGAAGACGGCGGCCACTACCTGTGAGTTCAAGTCCATC 1221  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1222 TACATGGCCCAAGAGCCCGTGCAGCTGCCGCTACTACTACGTGACTCCAAAGCTGAC 1281  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1282 ATCACCTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 1341  
Qy 221 HisHisLeuPheLeu 225  
Db 1342 CACCACCTGTTCTCTG 1356

RESULT 35

US-10-433-640-16  
; Sequence 16, Application US/10433640  
; Publication No. US20040115792A1  
; GENERAL INFORMATION:  
; APPLICANT: Lichtenberg-Frate, Hella  
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF  
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS  
; FILE REFERENCE: 1487/3  
; CURRENT APPLICATION NUMBER: US/10/433,640  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: PCT/BP01/14610  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: DE 10061872.3  
; PRIOR FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 16  
; LENGTH: 4692  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: sequecne of vector pDsRed1-N1  
US-10-433-640-16

Alignment Scores:

Pred. No.: 3.23e-141 Length: 4692  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) \* US-10-433-640-16 (1-4692)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 682 GTGGCTCTCTCCAAGAACGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 741  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 742 ACCGTGAACGGCCACGAGTTTCAGATCGAGGCGAGGGCGAGGGCCCTTACGAGGGC 801  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

Db 802 CACAACAACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGCTGGACATC 861  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 862 CTGTCCCCCAGTTCACAGTACCGGCTCCAAGGTGTACGTGAAGCACCCCGCAGATCCCC 921  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTTPGluArgValMetAsnPheGlu 100  
Db 922 GACTCAAGAAGCTGTCTTCTCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTGAG 981  
Qy 101 AspglyglyValValThrValThrGlnAspSerSerLeuGlnAspglyCysPheIleTyr 120  
Db 982 GACGGCGCGGTGGTGACCGGTGACCAAGACTCCTCCCTGCAAGACGGCTGCTCATCTAC 1041  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspglyProValMetGlnLysLysThr 140  
Db 1042 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCGTAAATGCAAGAAAGACC 1101  
Qy 141 MetGlyTTPGluAlaSerThrGluArgLeuTyrProArgAspglyValLeuLysGlyGlu 160  
Db 1102 ATGGGCTGGAGGCTCTCCACCAAGCCCTGTACCCCCCGCACGGCGCTGTAAGGGCGAG 1161  
Qy 161 IleHisLysAlaLeuLysLeuLysAspglyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 1162 ATCCACAAGGCCCTGAAGCTGAAGACGGCGCCACTACTGTGTGAGTTCAAGTTCATC 1221  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 1222 TACATGGCCCAAGAGCCCGTGCAGCTGCCGCTACTACTAGTGAATCCAAAGCTGAC 1281  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 1282 ATCACCTCCCAACAAGAGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCCGC 1341  
Qy 221 HisHisLeuPheLeu 225  
Db 1342 CACCACCTGTTCTCTG 1356

RESULT 36

US-11-006-076-29  
; Sequence 29, Application US/11006076  
; Publication No. US20050181506A1  
; GENERAL INFORMATION:  
; APPLICANT: Perkins, Edward  
; APPLICANT: Perez, Carl  
; APPLICANT: Lindenbaum, Michael  
; APPLICANT: Greene, Amy  
; APPLICANT: Leung, Josephine  
; APPLICANT: Fleming, Elena  
; APPLICANT: Stewart, Sandra  
; APPLICANT: Shellard, Joan  
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
; FILE REFERENCE: 17084-022002/420B  
; CURRENT APPLICATION NUMBER: US/11/006,076  
; CURRENT FILING DATE: 2004-12-06  
; PRIOR APPLICATION NUMBER: 60/294,758  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/366,891  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 4692  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pDsRed1-N1 plasmid from Clontech  
US-11-006-076-29

Alignment Scores: 3.23e-141 Length: 4692  
Pred. No.: 1210.00 Matches: 224  
Score:



Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-006-076-29 (1-4692)

```
QY      1 MetArgSerSerLyAsnValIleLygLuPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      682 GTGGCTCTCTCCAAGAACGTGATCAAGAGATTGATGCGCTTCAAGGTCCGCAATGAGGGC 741

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||
DB      742 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGAGGGCGAGGGCCGCCCTACGAGGGC 801

QY      41 HisAsnThrValLyLeuLyValThrLyGlyGlyProLeuProPheAlaTrpAspIle 60
      |||
DB      802 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCCCTTGGCTGGGACATC 861

QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80
      |||
DB      862 CTGTCCCCCAGTTCCAGTAGACGGCTCCAGGTGTAAGTGAAGCACCCCGCGACATCCCC 921

QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||
DB      922 GACTACAAAGAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTGAG 981

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
      |||
DB      982 GACGGCGCGTGTGACCGTGACCCAGACTCTCTCTGACGAGACGGCTGCTTCAATCTAC 1041

QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
      |||
DB      1042 AAGGTGAAGTTGATCGCGGTGAATTCCCTCCGACGGCCCCGTAATGCAAGAGAAGACC 1101

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      |||
DB      1102 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 1161

QY      161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180
      |||
DB      1162 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGAGTTCAAGTCCATC 1221

QY      181 TyrMetAlaLyLySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      |||
DB      1222 TACATGGCCAAAGAGCCCGTGACAGTCCCGGCTACTACTACTGAGTCCAAAGCTGAC 1281

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      |||
DB      1282 ATCACTCCCAACAAGAGACTACACCATGCTGAGGAGTACGAGCGCACCGAGGGCGC 1341

QY      221 HisHisLeuPheLeu 225
      |||
DB      1342 CACCACTGTTCCTG 1356
```

RESULT 37  
US-10-169-050-46

; Sequence 46, Application US/10169050  
; Publication No. US20050071891A1  
; GENERAL INFORMATION:  
; APPLICANT: THRESHER, RON  
; APPLICANT: HINDS, LYN  
; APPLICANT: HARDY, CHRIS  
; APPLICANT: WHYARD, STEVE  
; APPLICANT: VIGNARAJAN, SOMA  
; APPLICANT: GREWE, PETER MARTIN  
; APPLICANT: PATIL, JAWAHAR  
; TITLE OF INVENTION: REPRESSIBLE STERILITY OF ANIMALS  
; FILE REFERENCE: 4050.001500  
; CURRENT APPLICATION NUMBER: US/10/169,050  
; PRIOR FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: PCT/AU00/0  
; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: PQ4884  
; PRIOR FILING DATE: 1999-12-24  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 46  
; LENGTH: 5436  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pSFM6  
US-10-169-050-46

Alignment Scores:  
Pred. No.: 3.97e-141 Length: 5436  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-169-050-46 (1-5436)

```
QY      1 MetArgSerSerLyAsnValIleLygLuPheMetArgPheLyValArgMetGluGly 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      1426 GTGGCTCTCTCCAAGAACGTGATCAAGAGATTGATGCGCTTCAAGGTCCGCAATGAGGGC 1485

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||
DB      1486 ACCGTGAACGGCCACGAGTTGAGATCGAGGCGAGGGCGAGGGCCGCCCTACGAGGGC 1545

QY      41 HisAsnThrValLyLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60
      |||
DB      1546 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCCCTTGGCTGGGACATC 1605

QY      61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80
      |||
DB      1606 CTGTCCCCCAGTTCCAGTAGACGGCTCCAGGTGTAAGTGAAGCACCCCGCGACATCCC 1665

QY      81 AspTyrLyLyLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||
DB      1666 GACTACAGAAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAATTGAG 1725

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
      |||
DB      1726 GACGGCGCGTGTGACCGTGACCCAGACTCTCTCTGACGAGCGGCTGCTCATCTAC 1785

QY      121 LySValLySPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
      |||
DB      1786 AAGGTGAAGTTGATCGCGGTGAATTCCCTCCGACGGCCCCGTAATGCAAGAGAACC 1845

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      |||
DB      1846 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 1905

QY      161 IleHisLyAlaLeuLyLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180
      |||
DB      1906 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGAGTTCAAGTCCATC 1965

QY      181 TyrMetAlaLyLySProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
      |||
DB      1966 TACATGGCCAAAGAGCCCGTGACAGTCCCGGCTACTACTACTGAGTCCAAAGCTGAC 2025

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      |||
DB      2026 ATCACTCCCAACAAGAGACTACACCATGCTGAGGAGTACGAGCGCACCGAGGGCGCC 2085

QY      221 HisHisLeuPheLeu 225
      |||
DB      2086 CACCACTGTTCCTG 2100
```

RESULT 38

US-10-001-189-45  
; Sequence 45, Application US/10001189



```
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 6984
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: BXP3-DsRed-orf sequence
US-10-001-189-45

Alignment Scores:
Pred. No.:      5.62e-141      Length:      6984
Score:          1210.00      Matches:      224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches:  0
Query Match:    99.67%      Indels:      0
DB:             5           Gaps:          0

US-10-006-922A-12 (1-225) x US-10-001-189-45 (1-6984)

QY      1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20
Db      2974 GTGCGCTCCTCCAGAAGCTGTCATCAAGAGATTGCGCTTCAAGGTGCGCATGAGGCG 3033

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      3034 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 3093

QY      41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60
Db      3094 CACAACACCGGTGAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTTGGGACATC 3153

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
Db      3154 CTGTCCCCCAGTTCAGTAGTACGGCTCCAAGGTGTACGTGAAGCACCCCGCAGATCCCC 3213

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
Db      3214 GACTACAAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACCTTCGAG 3273

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db      3274 GACGGCGGCGTGTGACCGTGACCCAGACTCTCCCTGACAGACGGCGCTTCATCTAC 3333

QY      121 LysValLySPhetIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140
Db      3334 AAGGTGAAGTTGATCGGCGGTGAACCTCCCTCCGACGGCCCCGTAAATGACAGAAGACC 3393

QY      141 MetGlyTyrTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLy 160
Db      3394 ATGGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGGACCGGCGGTGTGAAGGGCGAG 3453

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180
Db      3454 ATCCACAAGGCGCTGAAGCTGAAGGACGGCGGCACTACCTGTGTGAGTTCAAGTTCATC 3513

QY      181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200
```

```
Db      3514 TACATGGCCAGAAGACCCGTGAGCTGCCCGGCTACTACTACTAGTGGACTCCAGCTGAC 3573

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db      3574 ATCACTCCCAACAACGAGGACTACCACTATCGTGAGCAGTACGACGCCACCGAGGGCCGC 3633

QY      221 HisHisLeuPheLeu 225
Db      3634 CACCACCTGTTCCTG 3648

RESULT 39
US-10-169-050-20/c
; Sequence 20, Application US/10169050
; Publication No. US20050071891A1
; GENERAL INFORMATION:
; APPLICANT: THRESHER, RON
; APPLICANT: HINDS, LYN
; APPLICANT: HARDY, CHRIS
; APPLICANT: WHYARD, STEVE
; APPLICANT: VIGNARAJAN, SOMA
; APPLICANT: GREWE, PETER MARTIN
; APPLICANT: PATIL, JAWAHAR
; TITLE OF INVENTION: REPRESSIBLE STERILITY OF ANIMALS
; FILE REFERENCE: 4050.001500
; CURRENT APPLICATION NUMBER: US/10/169,050
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: PCT/AU00/0
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: P04884
; PRIOR FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 7910
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pBIT(dhsp) -RFP-ohoxDS/BH
US-10-169-050-20

Alignment Scores:
Pred. No.:      6.69e-141      Length:      7910
Score:          1210.00      Matches:      224
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.56%      Mismatches:  0
Query Match:    99.67%      Indels:      0
DB:             9           Gaps:          0

US-10-006-922A-12 (1-225) x US-10-169-050-20 (1-7910)

QY      1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLySValArgMetGluGly 20
Db      2705 GTGCGCTCCTCCAAGAAGCTGATCAAGAGTTGCGCTTCAAGGTGCGCATGAGGGC 2646

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db      2645 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 2586

QY      41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60
Db      2585 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGGCCCCCTGCCCTTGGGACATC 2526

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
Db      2525 CTGTCCCCCAGTTCAGTAGCGGTCCAAAGGTGATGAGCAACCCCGACATCCCC 2466

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
Db      2465 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACCTTCGAG 2406

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
```



Db 2405 GACGGCGCGGTGACCGTGACCAGACTCTCTCCCTGCAGAGACGGCTGCTCATCTAC 2346  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140  
Db 2345 AAGGTGAAGTTCAATCGCGCTGAACCTCCCTCCGACGGCCCGTAATGCAAGAAAGACC 2286  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 2285 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGCTGTGAAGGGCGAG 2226  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 2225 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGGCCCACTACTGAGTGAAGTCCATC 2166  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 2165 TACATGGCCAAAGAGCCCGTGCAAGCTGCCCGCTACTACTAGTGAGACTCCAAAGCTGAC 2106  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnIntTyrGluArgThrGluGlyArg 220  
Db 2105 ATCACTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 2046  
Qy 221 HisHisLeuPheLeu 225  
Db 2045 CACCACTGTCTCTG 2031  
RESULT 40  
US-10-471-065-20  
; Sequence 20, Application US/10471065  
; Publication No. US20040197855A1  
; GENERAL INFORMATION:  
; APPLICANT: Prof. Dr. Wiesmuller, Lisa  
; TITLE OF INVENTION: Test system for the determination of genotoxicities  
; FILE REFERENCE: P59532  
; CURRENT APPLICATION NUMBER: US/10/471, 065  
; CURRENT FILING DATE: 2003-09-04  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 9320  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Description of the artificial sequence: Plasmid  
; OTHER INFORMATION: p5-Puro-CMV-(N'-EGFP)-(EGFP-EJ)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1592)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (3374)..(3392)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (5527)..(9320)  
; OTHER INFORMATION: Retroviral vector p5NM  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1617)..(2216)  
; OTHER INFORMATION: Puromycin resistance gene from pRetron (Clontech,  
; OTHER INFORMATION: Palo Alto, CA, USA)  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (2267)..(2848)  
; OTHER INFORMATION: CMV promoter from pEGFP-N1 (Clontech, Palo Alto,  
; OTHER INFORMATION: CA, USA)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (2906)..(3348)  
; OTHER INFORMATION: N'-EGFP, derived from EGFP from pEGFP-N1  
; OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)  
; FEATURE:

; NAME/KEY: promoter  
; LOCATION: (3411)..(3992)  
; OTHER INFORMATION: CMV promoter from pDsRed1-N1 (Clontech, Palo Alto,  
; OTHER INFORMATION: CA, USA)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (4038)..(4718)  
; OTHER INFORMATION: Red from pDsRed1-N1 (Clontech, Palo Alto, CA, USA)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (4766)..(5508)  
; OTHER INFORMATION: EGFP-EJ, derived from EGFP from pEGFP-N1  
; OTHER INFORMATION: (Clontech, Palo Alto, CA, USA)  
US-10-471-065-20  
Alignment Scores:  
Pred. No.: 8,4e-141 Length: 9320  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-471-065-20 (1-9320)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 4041 GTGCGCTCTCCAGAAGAGTCATCAAGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 4100  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 4101 ACCGTGAACGCCCAAGATTGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 4160  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 4161 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTGCGCTGGAGCATC 4220  
Qy 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 4221 CTGTCCCCCAGTTCCAGTACGAGCTCCAAAGGTGTAAGTGAACACCCCGCAGATCCCC 4280  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 4281 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGACTTCGAG 4340  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 4341 GACGGCGCGGTGTGACCGTGAACCCAGACTCTCTCCGACAGAGCGCTGCTCATCTAC 4400  
Qy 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 4401 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCCGAGCGGCCGTATGCAAGAAAGACC 4460  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 4461 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGCTGTGAAGGGCGAG 4520  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 4521 ATCCACAAGGCCCTGAAGCTGAAGAGCGCGGCCACTACTGAGTGAAGTCCAAAGTCCATC 4580  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 4581 TACATGGCCAAAGAGCCCGTGACAGCTGCCCGCTACTACTAGTGAGACTCCAAAGTGGAC 4640  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGlnIntTyrGluArgThrGluGlyArg 220  
Db 4641 ATCACTCCCAACAAGAGACTACACCATCGTGAAGCAGTACGAGCGCACCGAGGGCCGC 4700  
Qy 221 HisHisLeuPheLeu 225  
Db 4701 CACCACTGTCTCTG 4715







Db 5119 GTGCGCTCCTCCAGAACGTATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGGC 5178  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 5179 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC 5238  
QY 41 HisAsnThrValIleuLeuLysValThrIleGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 5239 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTCGCTCGCTGGGACATC 5298  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 5299 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTATCGTGAAGCACCCCGCCGACATCCCC 5358  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 5359 GACTACAAAGAGCTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 5418  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 5419 GACGGCGCGGTGTGACCGTGAACCCAGACTCTCCCTGACAGACGGCTGTTCATCTAC 5478  
QY 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
Db 5479 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCCGTATAGCAGAAAGAAC 5538  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 5539 ATGGGCTGGAGGCGCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 5598  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 5599 ATCCACAAGGCCCTGAAGCTGAAGGAGCGGCGCCACTACTGCTGAGTTCAGTTCATC 5658  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 5659 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACGTGACTCCAAAGCTGGAC 5718  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 5719 ATCACTCCCAACAAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCGGC 5778  
QY 221 HisHisLeuPheLeu 225  
Db 5779 CACCACTGTCTCTG 5793

RESULT 43  
US-10-609-019-2  
; Sequence 2, Application US/10609019  
; Publication No. US20040197910A1  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Richard K.  
; APPLICANT: Cadd, Gary G.  
; APPLICANT: Fioretti, William C.  
; APPLICANT: DeBoer, Kenneth F.  
; TITLE OF INVENTION: Gene Regulation in Transgenic Animals Using a Transposon-Based  
; TITLE OF INVENTION: Vector  
; FILE REFERENCE: 51687-0101 (51687-287015)  
; CURRENT APPLICATION NUMBER: US/10/609,019  
; CURRENT FILING DATE: 2003-06-26  
; PRIOR APPLICATION NUMBER: US 60/392,415  
; PRIOR FILING DATE: 2002-06-26  
; PRIOR APPLICATION NUMBER: US 60/441,392  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: US 60/441,377  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: US 60/441,502  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: US 60/441,405  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: US 60/441,447  
; PRIOR FILING DATE: 2003-01-21  
; PRIOR APPLICATION NUMBER: US 60/441,381

; PRIOR FILING DATE: 2003-01-21  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 10263  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-609-019-2

Alignment Scores:  
Pred. No.: 9,61e-141 Length: 10263  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 8 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-609-019-2 (1-10263)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 5725 GTGCGCTCCTCCAGAAAGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 5784  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 5785 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCGCCCTTACGAGGGC 5844  
QY 41 HisAsnThrValIleuLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 5845 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCGCTGGGACATC 5904  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 5905 CTGTCCCCCAGTTCCAGTACGGCTCCAAAGGTATCGTGAAGCACCCCGCCGACATCCCC 5964  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 5965 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 6024  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 6025 GACGGCGCGGTGTGACCGTGAACCCAGACTCTCTCCCTGACGAGCGGCTGCTGACTTAC 6084  
QY 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
Db 6085 AAGGTGAAGTTCATCGGCGTGAACCTTCCCTCCGACGGCCCCGTATGACAGAAAGAAC 6144  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLysGlyGlu 160  
Db 6145 ATGGGCTGGAGGCGCTTCACCGAGCGCCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 6204  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 6205 ATCCACAAGGCCCTGAAGCTGAAGGAGCGGCGCCACTACTGCTGAGTTCAAGTTCATC 6264  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
Db 6265 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGCTACTACTACGTGACTCCAAAGCTGGAC 6324  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 6325 ATCACTCCCAACAAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGGCGGC 6384  
QY 221 HisHisLeuPheLeu 225  
Db 6385 CACCACTGTCTCTG 6399

RESULT 44  
US-10-315-920-3  
; Sequence 3, Application US/10315920  
; Publication No. US20030175809A1



; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-3

Alignment Scores:
Pred. No.: 5.22e-142 Length: 678
Score: 1207.00 Matches: 223
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.11% Mismatches: 1
Query Match: 99.42% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-315-920-3 (1-678)

QY 1 MetArgSerSerLybAsnValIleLybGluPheMetArgPheLybValArgMetGluGly 20
DB 1 ATGCGCTCTCCAGAAAGCTCATCAAGAGATTCAATGCGCTTCAAGGTGGCATGAGGGC 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
DB 61 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCGCCCTTACGAGGGC 120
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
DB 121 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCGCTTGGAGCATC 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
DB 181 CTGTCCCCCAGTTCAGTAGCGGCTCCAGAGTGTACGTGAAGCACCCCCGCATCCCC 240
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTyrGluArgValMetAsnPheGlu 100
DB 241 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
DB 301 GACGGCGGCGTGGGCACTGTGACCCAGACTCTCTCTGACGAGCGGCTGCTCATCTAC 360
QY 121 LysValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLysThr 140
DB 361 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCGACGGCCCCGTGATGAGAAGAACCC 420
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyL 160
DB 421 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAG 480
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB 481 ATCCACAGAAGCGCTGAAGCTGAAGGACGGCGGCCCACTACTGTGTGAGTTCAAGTCCATC 540
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCGCGGCTACTACTACGTGACCAAGCTTGAC 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220

DB 601 ATCACCTCCACAACGAGAGCTACCATCGTGAGCAGTACGAGCGACACCGAGGGCCGC 660
QY 221 HisHisLeuPheLeu 225
DB 661 CACCACCTGTTCCTG 675

RESULT 45

US-10-742-828-4
; Sequence 4, Application US/10742828
; Publication No. US20040157294A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich, Gerhard
; APPLICANT: Huynh, G191
; TITLE OF INVENTION: TRANSGENIC SCREEN AND METHOD FOR SCREENING MODULATORS OF BRAIN-DEF
; FILE REFERENCE: US 1353/03 (VA)
; CURRENT APPLICATION NUMBER: US/10/742,828
; PRIOR FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 10/306,737
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 4
; LENGTH: 7508
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Unsure
; LOCATION:
; OTHER INFORMATION: Synthesized
US-10-742-828-4

Alignment Scores:
Pred. No.: 4.76e-140 Length: 7508
Score: 1203.00 Matches: 222
Percent Similarity: 99.56% Conservative: 2
Best Local Similarity: 98.67% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-742-828-4 (1-7508)

QY 1 MetArgSerSerLybAsnValIleLybGluPheMetArgPheLybValArgMetGluGly 20
DB 2123 GTGCGCTCTCCAGAAAGCTCATCAAGAGTTCAATGCGCTTCAAGGTGGCATGAGGGC 2182
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrgLugly 40
DB 2183 ACCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGGCGCCCTTACGAGGGC 2242
QY 41 HisAsnThrValLybLeuLybValThrLybGlyGlyProLeuProPheAlaTrpAspIle 60
DB 2243 CACAACACCGTGAAGCTGAAGGTGACCAAGGCGGCCCCCTGCGCTTGGAGCATC 2302
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
DB 2303 CTGTCCCCCAGTTCAGTAGCGGCTCCAGAGTGTACGTGAAGCACCCCCGCATCCCC 2362
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTyrGluArgValMetAsnPheGlu 100
DB 2363 GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 2422
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120
DB 2423 GACGGCGGCGTGGGCACTGTGACCCAGACTCTCTCTGACGAGCGGCTGCTTATCTAC 2482
QY 121 LysValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLysThr 140
DB 2483 AAGGTGAAGTTCAATCGCGCTGAACCTTCCCTCGACGGCCCCGTGATGAGAAGAACCC 2542
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLybGlyL 160
DB 2543 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGGCTGTGAAGGGCGAG 2602







|||||  
Db 61 CACGAGTTGAAATAGAAGCGGAGAGGAGGAGGCCATTACGAAGGCCACAATACCGTA 120  
Qy 45 LysLeuLysValThrLysGlyGlyProLeuProPheAlaTpaAspIleLeuSerProGln 64  
Db 121 AAGCTTAAGGTAAACCAAGGGGGGACCCTTTGCCATTGCTTGGAATATTTGTACACCAAA 180  
Qy 65 PheGlnTyrglySerLysValTyrrValLysHisProAlaAspIleProAspTyrrLysLys 84  
Db 181 TTTCAGTATGGAAGCAAGGTATATGTCAAGCAACCTGCCGACATACGACTATATAAAAG 240  
Qy 85 LeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyVal 104  
Db 241 CTGTCAATTCCTGAAGGATTTAAATGGAAAGGGTCATGAACCTTGAAAGACGGTGGCGTC 300  
Qy 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrrLysValLysPhe 124  
Db 301 GTTACTGTAAACCAAGGATTCAGATTGCAAGATGGCTGTTTCATCTAAGTCAAGTTC 360  
Qy 125 IleGlyValAlaAsnPheProSerSerLeuGlnAspGlyProValMetGlnLysLysThrMetGlyTrpGlu 144  
Db 361 ATTGGCGTGAACCTTTCCTCCGATGGAACCTGTTATGCCAAAAGAAAGCAATGGCGTGGAA 420  
Qy 145 AlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLysGlyGluIleHisLysAla 164  
Db 421 GCCAGCACTGAGCGTTGTATCCTCGTGAAGCGGTGTGAAAGAGAGATTCAATAAGGCT 480  
Qy 165 LeuLysLeuLysAspGlyGlyHisTyrrLeuValGluPheLysSerIleTyrrMetAlaLys 184  
Db 481 CTGAAGCTGAAGACGGGTGTCATTACCTAGTTGAATTCAAAAGTATTTAATGGAAG 540  
Qy 185 LysProValGlnLeuProGlyTyrrTyrrTyrrValAspSerLysLeuAspIleThrSerHis 204  
Db 541 AAGCCTGTGACGTAACCAAGGTAATACTGTAAGTCCAAACTGGAATAACAAGCCAC 600  
Qy 205 AsnGluAspTyrrThrIleValGluGlnTyrrGluArgThrGluGluArgHisIleLeuPhe 224  
Db 601 AACGAAGACTATACAATCGTTGAGCAGTATGAAGAACGAGGAGCGCACCATCTGTTC 660  
Qy 225 Leu 225  
Db 661 CTT 663

RESULT 48  
US-10-314-936-1  
; Sequence 1, Application US/10314936  
; Publication No. US20040110225A1  
; GENERAL INFORMATION:  
; APPLICANT: Gibbs, Patrick D.L.  
; APPLICANT: Carter, Robert W.  
; APPLICANT: Schmale, Michael C.  
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
; FILE REFERENCE: 638.004  
; CURRENT APPLICATION NUMBER: US/10/314,936  
; CURRENT FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: mutant red fluorescent protein  
; NAME/KEY: CDS  
; LOCATION: (1)..(711)  
; OTHER INFORMATION:  
US-10-314-936-1

Alignment Scores:  
Pred. No.: 1.37e-140 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2

Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-314-936-1 (1-711)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGAGTTGTTCCAAGAATGTATCAAGAGTTTCATGAGGTTTAAGTTCGTATGAAGGA 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGluGlyGluGlyArgProTyrrGluGly 40  
Db 61 ACGGTCAATGGGCACGAGTTGAAATAGAAGCGGAAGGAGGAGGCCATACGAAGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTpaAspIle 60  
Db 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGGGACCCTTTGCCATTGCTTGGAATAT 180  
Qy 61 LeuSerProGlnPheGlnTyrglySerLysValTyrrValLysHisProAlaAspIlePro 80  
Db 181 TTGTACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAAGCATCCTGCCACATACCA 240  
Qy 81 AspTyrrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATATAAAAGCTGTCAATTCCTGAAGATTTAAATGGAAAGGTCATGAACCTTGA 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrr 120  
Db 301 GACGGTGGCGTCGTTACTGTAAACCAAGATTCAGTTGCAAGATGGCTGTTTCATCTAC 360  
Qy 121 LysValLysPheIleGlyValAlaAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTCAAGTTCAATTGGCGTGAACCTTCTTGATGGAACCTGTTATGCAAAAGAAACA 420  
Qy 141 MetGlyTyrrGluAlaSerThrGluArgLeuTyrrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGGAAGCCAGCACTGAGCGTTGTATCCTCGTGAAGCGGTGTGAAAGAGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrrLeuValGluPheLysSerIle 180  
Db 481 ATTCATTAAGCTCTGAAGTTGAAGACGGGTGTCATTACCTAGTTGAATTCAAAACTATT 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrrTyrrTyrrValAspSerLysLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGACGTAACCAAGGTAATACTGTAAGTCCAAACTGAT 600  
Qy 201 IleThrSerHisAsnGluAspTyrrThrIleValGluGlnTyrrGluArgThrGluGluArg 220  
Db 601 ATAACAAGCCACAACAAGAATAATCGTTGAGCAGTATGAAGAAGAACGAGGAGCGC 660  
Qy 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

RESULT 49  
US-10-314-936-3  
; Sequence 3, Application US/10314936  
; Publication No. US20040110225A1  
; GENERAL INFORMATION:  
; APPLICANT: Gibbs, Patrick D.L.  
; APPLICANT: Carter, Robert W.  
; APPLICANT: Schmale, Michael C.  
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
; FILE REFERENCE: 638.004  
; CURRENT APPLICATION NUMBER: US/10/314,936  
; CURRENT FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:



OTHER INFORMATION: mutant red fluorescent protein  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(711)  
OTHER INFORMATION:  
US-10-314-936-3

Alignment Scores:  
Pred. No.: 1.37e-140 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-314-936-3 (1-711)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAAGAATGTTATCAAGAGTTTCATGAGTTTAAGTTCGTATGGAAGCA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCAAGAGTTGAATAAGAGCGAAGGAGGAGGCCATACGAAGCC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGAGCCTTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAAGTAAAGCAAGGATATATGCAAGCATCTGCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAGCTGTCAATTCGTAAGAGATTAAATGGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACCGTGGCGTCTGTAATCTGTAACCAAGATTCAGTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTTGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGCGTGGGAAGCCAGCACGTAGCGTTGTATCCTCGTATGCGCGTGTGAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCAATAAGCTCTGAAGTTGAAGACGGTGTCTATTAACCTTAATCAAAACTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCAAGAAGCCTGTGCAAGTACCAAGGTAATACTATGTAATCCCAACTGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAGCCACAACAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCATCTGTTCCCTT 675

RESULT 50  
US-11-021-014-1  
Sequence 1, Application US/11021014  
Publication No. US20050100954A1  
GENERAL INFORMATION:  
APPLICANT: Gibbs, Patrick D.L.  
APPLICANT: Carter, Robert W.  
APPLICANT: Schmale, Michael C.

TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
FILE REFERENCE: 638.004  
CURRENT APPLICATION NUMBER: US/11/021,014  
CURRENT FILING DATE: 2004-12-23  
PRIOR APPLICATION NUMBER: US/10/314,936  
PRIOR FILING DATE: 2002-12-09  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 711  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: mutant red fluorescent protein  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(711)  
OTHER INFORMATION:  
US-11-021-014-1

Alignment Scores:  
Pred. No.: 1.37e-140 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-021-014-1 (1-711)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGAGTTGTTCCAAGAATGTTATCAAGAGTTTCATGAGTTTAAGTTCGTATGGAAGCA 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACGGTCAATGGGCAAGAGTTGAATAAGAGCGAAGGAGGAGGCCATACGAAGCC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 121 CACAATACCGTAAGCTTAAGGTAAACCAAGGGGAGCCTTTGCCATTGCTGGGATATT 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 TTGTCAACCAATTTCAAGTAAAGCAAGGATATATGCAAGCATCTGCGACATACCA 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTATAAAAGCTGTCAATTCGTAAGAGATTAAATGGGAAAGGTCATGAACCTTGAA 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACCGTGGCGTCTGTAATCTGTAACCAAGATTCAGTTGCAAGATGGCTGTTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTCAAGTTCAATGGCGTGAACCTTCTTGATGACCTGTTATGCAAAAGAGACA 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGCGTGGGAAGCCAGCACTAGCGTTGTATCCTCGTATGCGCGTGTGAAGGAGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATTCAATAAGCTCTGAAGTTGAAGACGGTGTCTATTAACCTTAATCAAAACTATT 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCAAGAAGCCTGTGCAAGTACCAAGGTAATACTATGTAATCCCAACTGAT 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATAACAGCCACAACAAGACTATACATCGTTGAGCAGTATGAAGAACCAGAGGAGCGC 660



OY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

RESULT 51

US-11-021-014-3  
; Sequence 3, Application US/11021014  
; Publication No. US20050100954A1  
; GENERAL INFORMATION:  
; APPLICANT: Gibbs, Patrick D.L.  
; APPLICANT: Carter, Robert W.  
; APPLICANT: Schmale, Michael C.  
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES  
; FILE REFERENCE: 638.004  
; CURRENT APPLICATION NUMBER: US/11/021,014  
; CURRENT FILING DATE: 2004-12-23  
; PRIOR APPLICATION NUMBER: US/10/314,936  
; PRIOR FILING DATE: 2002-12-09  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: mutant red fluorescent protein  
; NAME/KEY: CDS  
; LOCATION: (1)..(711)  
; OTHER INFORMATION:  
US-11-021-014-3

Alignment Scores:  
Pred. No.: 1.37e-140 Length: 711  
Score: 1196.00 Matches: 221  
Percent Similarity: 99.11% Conservative: 2  
Best Local Similarity: 98.22% Mismatches: 2  
Query Match: 98.52% Indels: 0  
DB: 10 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-021-014-3 (1-711)

OY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGAGTTGTTCCAAGAATGTTATCAAGAGTTCATGAGGTTTAAGTTCGTATGGAAGCA 60  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACGGTCAATGGGCAAGAGTTTGAATAGAGCGAAGAGGAGGACCATACGAAGGC 120  
OY 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAATACCGTAAGAAGCTTAAGGTAACCAAGGGGGAACCTTGGCAATTGCTTGGATATT 180  
OY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 TTGTCACCAACAATTTCAAGTATGGAAGCAAGGTATATGTCAAGCATCTGCCGACATACCA 240  
OY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTATAAAGCTGTCATTTCTGGAAGATTAAATGGAAAGGTCATGAACCTTGAA 300  
OY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGTGGCGTCTACTGTAAACCAAGATCCAGTTTGACAGATGGCTTTCATCTAC 360  
OY 121 LySValLySLeuPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTCAAGTTCAATTGGCGGTGAACCTTCTTCGTATGACCTGTATGCAAAAGAGACA 420  
OY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGLyGlu 160

Db 421 ATGGCTGGGAAGCCAGCACTGAGCGTTTGTATCCTCGTGAATGCGGTGTGAAGAGAG 480  
OY 161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeuIle 180  
Db 481 ATTCAATAAGCTCTGAAGTTGAAGAAGCGTGGTTCATTACCTAGTTGAATTCAAACTATT 540  
OY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCAAGAAGCCTGTGCGCTACCAAGGTAATGTTGACTCCAACTGAT 600  
OY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATAACAAGCCACACAACAAGACTATACATCGTTGAGCAGATGATAAAGAAGGAGCGC 660  
OY 221 HisHisLeuPheLeu 225  
Db 661 CACCATCTGTTCTT 675

RESULT 52

US-10-006-922-36  
; Sequence 36, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Terexikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/457,556  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/444,338  
; PRIOR FILING DATE: 1999-11-19  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma species  
US-10-006-922-36

Alignment Scores:  
Pred. No.: 5.47e-140 Length: 678  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-006-922-36 (1-678)

OY 1 MetArgSerSerLyAsnValIleLyGluPheMetArgPheLySValArgMetGluGly 20  
Db 1 ATGGCTCTCCGAGAAGCTGATCAACGAAGTTTCATGCGCTTCAAGTGCATGAGAGGC 60  
OY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAAGGCCACGAGTTCGATCGAGGCGAGGCGGCGCCCTACGAGAGGC 120  
OY 41 HisAsnThrValLySLeuLySValThrLySGLyGlyProLeuProPheAlaTrpAspIle 60



DB 121 CACAACACCGTGAAGTGAAGGTGACCAAGGCGGCCCCCTGCTCCCTTGGGACATC 180  
QY LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCAGTTCAGTACGCGCTCCAAGGTGTACGTGAAGCACCCCGCATATCCCC 240  
QY AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACAGAAGAGCTGCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300  
QY AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGCGCGCGTGGCGACCGGTGACCGACGACTCTCTCCCTGACAGACGCGCTGCTCATCTAC 360  
QY LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140  
DB 361 AAGGTGAAGTTCAATCGCGGTGAATCTCCCTCCGACGCGCCCGTGATGCAAGAGAACC 420  
QY MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCGACGCGGTGTGAAGGGCGAG 480  
QY 161 ILeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCCACAAGGCCCTGAAGTGAAGAGCGCGGCCACTACTGTGTGAGTTCAATCTCATC 540  
QY TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAGAAGACCCCTGCAAGCTGCGCGGCTACTACTACGTGACACCAAGCTGGAC 600  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTTCCACACGAGGACTACACCATCGTGAGAGTAAGAGCGCACCGAGGGCGGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACCTGTCTCTG 675

RESULT 53  
US-10-081-864-14  
; Sequence 14, Application US/10081864  
; Publication NO. US20030022287A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey  
; APPLICANT: Lukyanov, Konstantin  
; APPLICANT: Yanushevich, Yuriy  
; APPLICANT: Savitsky, Alexandr  
; APPLICANT: Pradkov, Arcady  
; TITLE OF INVENTION: No. US20030022287A1 Aggregating Fluorescent Proteins and  
; TITLE OF INVENTION: Methods for Using the Same  
; FILE REFERENCE: CLON-067  
; CURRENT APPLICATION NUMBER: US/10/081,864  
; CURRENT FILING DATE: 2002-06-19  
; PRIOR APPLICATION NUMBER: 10/006,922  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 60/270,983  
; PRIOR FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: non-aggregating mutant  
US-10-081-864-14

Alignment Scores:  
Pred. No.: 5.47e-140 Length: 678  
Score: 1191.00 Matches: 220  
Percent Similarity: 98.67% Conservative: 2  
Best Local Similarity: 97.78% Mismatches: 3  
Query Match: 98.11% Indels: 0

DB: 5 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-081-864-14 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCTCTCTCCGAGAAGCTCATACCGAGTTTCATGCGCTTCAAGGTGGCATGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 CACAACACCGTGAAGTGAAGGTGACCAAGGCGGCCCCCTGCTCCCTTGGGACATC 180  
QY LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCCCAGTTCAGTACGCGCTCCAAGGTGTACGTGAAGCACCCCGCATATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACAGAAGAGTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGCGCGCGTGGCGACCGGTGACCGAGACTCTCTCCCTGACAGACGCGCTGCTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysThr 140  
DB 361 AAGGTGAAGTTCAATCGCGGTGAATCTCCCTCCGACGCGCCCGTGATGCAAGAGAACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCGCGACGCGGTGTGAAGGGCGAG 480  
QY 161 ILeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCCACAAGGCCCTGAAGTGAAGAGCGCGGCCACTACTGTGTGAGTTCAATCTCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAGAAGACCCCTGCAAGCTGCGCGGCTACTACTACGTGACACCAAGCTGGAC 600  
QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTTCCACACGAGGACTACACCATCGTGAGAGTAAGAGCGCACCGAGGGCGGC 660  
QY 221 HisHisLeuPheLeu 225  
DB 661 CACCACCTGTCTCTG 675

RESULT 54  
US-10-315-920-5  
; Sequence 5, Application US/10315920  
; Publication No. US20030175809A1  
; GENERAL INFORMATION:  
; APPLICANT: Pradkov, Arcady Fedorovich  
; APPLICANT: Tersikh, Alexey  
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS  
; TITLE OF INVENTION: FOR THEIR USE  
; FILE REFERENCE: CLON-077CIP  
; CURRENT APPLICATION NUMBER: US/10/315,920  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/211,607  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: PCT/US01/19097  
; PRIOR FILING DATE: 2001-06-13  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA







Db 541 TACATGGCCAAAGAGCCCGTCAGCTGCCCGGCTACTACTAGTGAAGCCCAAGCTGAC 600  
QY 201 ILeThSerHisAnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIYArg 220  
Db 601 ATCACCTCCCAAGAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGCCCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTG 675  
RESULT 56  
US-10-081-864-13  
/ Sequence 13, Application US/10081864  
/ Publication No. US2003002287A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Lukyanov, Sergey  
/ APPLICANT: Lukyanov, Konstantin  
/ APPLICANT: Yanushevich, Yuriy  
/ APPLICANT: Savitsky, Alexander  
/ APPLICANT: Pradkov, Arcady  
/ TITLE OF INVENTION: No. US2003002287A1 Aggregating Fluorescent Proteins and  
/ TITLE OF INVENTION: Methods for Using the Same  
/ FILE REFERENCE: CLON-067  
/ CURRENT APPLICATION NUMBER: US/10/081,864  
/ CURRENT FILING DATE: 2002-06-19  
/ PRIOR APPLICATION NUMBER: 10/006,922  
/ PRIOR FILING DATE: 2001-12-04  
/ PRIOR APPLICATION NUMBER: 60/270,983  
/ PRIOR FILING DATE: 2001-02-21  
/ NUMBER OF SEQ ID NOS: 30  
/ SOFTWARE: FastSeq for windows Version 4.0  
/ SEQ ID NO 13  
/ LENGTH: 675  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: non-aggregating mutant  
US-10-081-864-13  
Alignment Scores:  
Pred. No.: 2,32e-139 Length: 675  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 5 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-081-864-13 (1-675)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
Db 1 ATGGCCTCCTCCGAGAACGTCAATCAACGAGTTCAATGCGCTTCAAGGTGCGCATGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGCG 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 CTGTCCCGCCCAAGTTCAGTACGGCTCAAGGTGATGAGCAACCCCGACATCCCC 240  
QY 81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAGAAGACTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGGCGAACCGTGAACCGAGACTCTCCCTGACAGACGGCTGCTTCACTAC 360

QY 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeuThr 140  
Db 361 AAGGTGAAGTTCATCGCGGTGAACCTTCCCTCCGACGCGCCCGGTGATGCAAGAAAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValIleuLySGlyGlu 160  
Db 421 ATGGGCTGGAGGCTTCCACCGAGCGCTGTACCCCCGCGACGCGGTGCTGAAGGGCGAG 480  
QY 161 ILeHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySLeu 180  
Db 481 ACCCAAGGCCCTGAAGCTGAAGGACGGCGGCCCACTACCTGTGAGTTCAAGTCCATC 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCCAAAGAGCCCGTCAGCTGCCGCTACTACTACGTGACGCCCAAGCTGAC 600  
QY 201 ILeThSerHisAnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIYArg 220  
Db 601 ATCACCTCCCAAGAGAGACTACACCATCGTGAGCAGTACGAGCGCACCGAGGCCCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCTG 675  
RESULT 57  
US-10-785-862-10  
/ Sequence 10, Application US/10785862  
/ Publication No. US20040261149A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Donald Danforth Plant Science Center  
/ APPLICANT: Fauquet, Claude M.  
/ APPLICANT: Padmanabhan, Chellappan  
/ APPLICANT: Ramachandran, Vanitharani  
/ TITLE OF INVENTION: siRNA-mediated inhibition of gene expression in plant cells  
/ FILE REFERENCE: C35621/104850  
/ CURRENT APPLICATION NUMBER: US/10/785,862  
/ CURRENT FILING DATE: 2004-02-24  
/ NUMBER OF SEQ ID NOS: 10  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 10  
/ LENGTH: 747  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: the dsred coding sequence  
US-10-785-862-10  
Alignment Scores:  
Pred. No.: 2,68e-139 Length: 747  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 8 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-785-862-10 (1-747)  
QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLySValArgMetGluGly 20  
Db 1 ATGGCCTCCTCCGAGAACGTCAATCAACGAGTTCAATGCGCTTCAAGGTGCGCATGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTCAGATCGAGGGCGAGGGCGGCCCTTACGAGGCG 120  
QY 41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCTTGGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 CTGTCCCGCCCAAGTTCAGTACGGCTCAAGGTGATGAGCAACCCCGACATCCCC 240







Query Match: 97.69% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-894-949-9 (1-4200)

```
OY      1 MetArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB      3267 ATGGCCTCCTCCGAGAACGTATCATCACCGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 3326
OY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB      3327 ACCGTGAACGGCCACGAGTTCCAGATCGAGGCGAGGGCGAGGGCCGCCCTTACGAGGGC 3386
OY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
DB      3387 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGGTGGACATC 3446
OY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
DB      3447 CTGTCCCCCAGTTCCAGTACGAGCTCCAGAGGTGATGAAGCACCCCGCGACATCCCC 3506
OY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
DB      3507 GACTACAAAGAGCTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCGAG 3566
OY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      3567 GACGGCGCGGTGGCGAACCCTGACCCAGACTCTCTCTGACAGACGGCTGTTCATCTAC 3626
OY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB      3627 AAGGTGAAGTTCATCGGCGTGAATTCCCTCCGACGGCCCCGTGATGACAGAAAGACC 3686
OY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB      3687 ATGGGCTGGAGGGCTCCACCGAGCGCCTGTACCCCGCGACGGCGTGTGAAGGGCGAG 3746
OY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB      3747 ACCCACAAAGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 3806
OY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB      3807 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTGAGCCCAAGCTGGAC 3866
OY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      3867 ATCACTCCCAACAAGAGACTACACCATCGTGAAGCACTACGAGCGCACCGAGGGCGGC 3926
OY      221 HisHisLeuPheLeu 225
DB      3927 CACCACCTGTCTCTG 3941

RESULT 60
US-10-894-949-8
; Sequence 8, Application US/10894949
; Publication No. US20050130919A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Zuoshang
; APPLICANT: Xia, Xugang
; TITLE OF INVENTION: REGULATABLE PROMOTERS FOR SYNTHESIS OF SMALL HAIRPIN RNA
; FILE REFERENCE: UMY-072
; CURRENT APPLICATION NUMBER: US/10/894,949
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/488,510
; PRIOR FILING DATE: 2003-07-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

; OTHER INFORMATION: expression construct  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 536, 2369  
; OTHER INFORMATION: n = A, T, C or G  
US-10-894-949-8

Alignment Scores:  
Pred. No.: 3.06e-138 Length: 4300  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-894-949-8 (1-4300)

```
OY      1 MetArgSerSerIysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
DB      1291 ATGGCCTCCTCCGAGAACGTATCATCACCGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC 1350
OY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB      1351 ACCGTGAACGGCCACGAGTTCCAGATCGAGGCGAGGGCGAGGGCCGCCCTTACGAGGGC 1410
OY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60
DB      1411 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGGTGGACATC 1470
OY      61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80
DB      1471 CTGTCCCCCAGTTCCAGTACGAGCTCCAGAGGTGATGAAGCACCCCGCGACATCCCC 1530
OY      81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100
DB      1531 GACTACAAAGAGCTGTCTTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCGAG 1590
OY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      1591 GACGGCGCGGTGGCGAACCCTGACCCAGACTCTCTCTGACAGACGGCTGTTCATCTAC 1650
OY      121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140
DB      1651 AAGGTGAAGTTCATCGGCGTGAATTCCCTCCGACGGCCCCGTGATGACAGAAAGACC 1710
OY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160
DB      1711 ATGGGCTGGAGGGCTCCACCGAGCGCCTGTACTACCCCGCGACGGCGTGTGAAGGGCGAG 1770
OY      161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
DB      1771 ACCCACAAAGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 1830
OY      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
DB      1831 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTGAGCGCAAGCTGGAC 1890
OY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      1891 ATCACTCCCAACAAGAGACTACACCATCGTGAAGCACTACGAGCGCACCGAGGGCGGC 1950
OY      221 HisHisLeuPheLeu 225
DB      1951 CACCACCTGTCTCTG 1965

RESULT 61
US-10-742-828-5
; Sequence 5, Application US/10742828
; Publication No. US20040157294A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich, Gerhard
; APPLICANT: Huynh, Gigi
; TITLE OF INVENTION: TRANSGENIC SCREEN AND METHOD FOR SCREENING MODULATORS OF BRAIN-DEF
```



```

; FILE REFERENCE: US 1353/03 (VA)
; CURRENT APPLICATION NUMBER: US/10/742,828
; CURRENT FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 10/306,737
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 5
; LENGTH: 7495
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Unsure
; LOCATION:
; OTHER INFORMATION: Synthesized
US-10-742-828-5

Alignment Scores:
Pred. No.:      6.64e-138      Length:      7495
Score:          1186.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%    Mismatches:  4
Query Match:     97.69%        Indels:       0
DB:              7            Gaps:          0

US-10-006-922A-12 (1-225) x US-10-742-828-5 (1-7495)

QY      1 MetArgSerSerLySAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20
DB      2123 ATGGCCTCTCCGAGAACGTATCATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGCG 2182

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGlyArgProTyrGluGly 40
DB      2183 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCGGCCCTTACGAGGGCG 2242

QY      41 HisAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60
DB      2243 CACAACACCGTGAAAGCTGAAGGTGACCAAGGCGGCCCTTGCTTCCCTGGGACATC 2302

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
DB      2303 CTGTCCCCCAGTTCAGATACGGCTCCAGAGGTGTACGTGAAGCACCCGCCGACATCCCC 2362

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB      2363 GACTACAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 2422

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      2423 GACGGCGGCGTGGCGAACCCTGATCCCTCTCCCTGACGAGCGGCTGCTTCACTAC 2482

QY      121 LysValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140
DB      2483 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGCCCCCGTATGCAAGAAAGACC 2542

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySgLyGlu 160
DB      2543 ATGGGCTGGGAGGCTCCACCGAGCGGCTGTACCCCGGACGGCGGTGTAAGGGCGAG 2602

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180
DB      2603 ACCCACAAGGCCCTGAAGCTGAAGAGACGGGGCCACTACCTGGTGAAGTTCAGTCTATC 2662

QY      181 TyrMetAlaLySlySProValGlnLeuProGluTyrTyrTyrValAspSerLySleuAsp 200
DB      2663 TACATGGCCACAAGAACCCCGTCAAGCTGCGCGGCTACTACTACGTGAGCGCCAAAGCTGGAC 2722

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      2723 ATCACTCTCCACAACGAGGACTACACCATCTGTGAGCAGTACGAGCGCACCGAGGGCGCGC 2782

QY      221 HisHisLeuPheLeu 225
DB      2783 CACCACCTGTCTCTG 2797
```

```

RESULT 62
US-10-510-363-5
; Sequence 5, Application US/10510363
; Publication No. US20050158724A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Richard A.
; APPLICANT: Yoder, Kristine E.
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT MODULATE A DNA REPAIR
; TITLE OF INVENTION: PATHWAY AND/OR RETRO-VIRAL INFECTIVITY, THE COMPOUNDS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SCHN-0033
; CURRENT APPLICATION NUMBER: US/10/510,363
; CURRENT FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/US03/10302
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,376
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 9731
; TYPE: DNA
; ORGANISM: Artificial
; OTHER INFORMATION: Synthesized retroviral vectors
US-10-510-363-5

Alignment Scores:
Pred. No.:      9.55e-138      Length:      9731
Score:          1186.00        Matches:      219
Percent Similarity: 98.22%      Conservative: 2
Best Local Similarity: 97.33%    Mismatches:  4
Query Match:     97.69%        Indels:       0
DB:              9            Gaps:          0

US-10-006-922A-12 (1-225) x US-10-510-363-5 (1-9731)

QY      1 MetArgSerSerLySAsnValIleLySgluPheMetArgPheLySValArgMetGluGly 20
DB      5496 ATGGCCTCTCCGAGAACGTATCATCACCGAGTTTCATGCGCTTCAAGGTGCGCATGGAGGCG 5555

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGlyGlyGlyArgProTyrGluGly 40
DB      5556 ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGGCGGCCCTTACGAGGGCG 5615

QY      41 HisAsnThrValLySLeuLySValThrLySgLyGlyProLeuProPheAlaTrpAspIle 60
DB      5616 CACAACACCGTGAAAGCTGAAGGTGACCAAGGCGGCCCTTGCTTCCCTGGGACATC 5675

QY      61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
DB      5676 CTGTCCCCCAGTTCAGATACGGCTCCAAAGGTGTACGTGAAGCACCCCGGACATCCCC 5735

QY      81 AspTyrLySlySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
DB      5736 GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAATTCGAG 5795

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      5796 GACGGCGGCGTGGCGAACCCTGATCCCTCCCGAGCGGCTGATGAATTCGAG 5855

QY      121 LysValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySlyThr 140
DB      5856 AAGGTGAAGTTTCATCGCGGTGAACCTTCCCTCCGACGGGCCCTGATGCAAGAAAGACC 5915

QY      141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySgLyGlu 160
DB      5916 ATGGGCTGGGAGGCTCCACCGAGCGGCTGTACCCCGGACGGCGGTGTAAGGGCGAG 5975

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180
DB      5976 ACCCAAGGCCCTGAAGCTGAAGAGACGGCGGCGCACTACTGCTGTGAAGTTCATC 6035
```



OY	181	TyrMetAlaIysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
			: :
Db	6036	TACATGGCCAGAGAAGCCCGTGACGTGCCCGCTACTACTACGTGACGCACAAGCTGGAC	6095
OY	201	IleThrSerHisasnGluAspTyrThrIleValGlnGlnTyrGluArgThrGlnGlyArg	220
Db	6096	ATCAACCTCCCAACAAGAGACTACACCATCGTGAGCAGTAGCAGCACCAAGGGCCGC	6155
OY	221	HisHisLeuPheLeu	225
Db	6156	CACCACCTGTTCTCG	6170

```

RESULT 63
US-10-510-363-6
; Sequence 6, Application US/10510363
; Publication No. US20050158724A1
; GENERAL INFORMATION:
; APPLICANT: Fishel, Richard A.
; APPLICANT: Yoder, Kristine E.
; TITLE OF INVENTION: METHODS OF IDENTIFYING COMPOUNDS THAT MODULATE A DNA REPAIR
; TITLE OF INVENTION: PATHWAY AND/OR RETRO-VIRAL INFECTIVITY, THE COMPOUNDS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: SCHN-0033
; CURRENT APPLICATION NUMBER: US/10/510,363
; CURRENT FILING DATE: 2004-10-05
; PRIOR APPLICATION NUMBER: PCT/US03/10302
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,376
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 9782
;
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthesized retroviral vectors
US-10-510-363-6

```

Alignment Scores:	
Pred. No.:	9.62e-138
Score:	1186.00
Percent Similarity:	98.22%
Best Local Similarity:	97.33%
Query Match:	97.69%
DB:	9
Length:	9782
Matches:	219
Conservative:	2
Mismatches:	4
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-10-510-363-6 (1-9782)

QY	1	MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly	20
		: :	
Db	13	ATGACCTCTCCGAGAACGTATCAACGAGTTCATGCGCTTCAAGGTGCGCATGAGAGGC	72
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	73	ACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGCGCCCTACGAGGGC	132
QY	41	HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	133	CACACACCCGTGAAGCTGAAGGTGACCAAGGGCGGGCCCTGCCCTTGCGCTGGACATC	192
QY	61	LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro	80
Db	193	CTGTCCCCCAGTTCACGATGAGGCTCCAAAGGTGACGTGAAGCACCCCGCCGACATCCCC	252
QY	81	AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu	100
Db	253	GACTACAGAAAGCTGTCTCTCCCGAGGGCTTCAAGTGGGAGCGCGGTGATGAACCTTCGAG	312
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	313	GACGGCGGCGTGGCGACCGGTGACCCAGACTCTCCCTGCAGAGACGGCTGTCATCTAC	372

QY	121	LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIlySlyeThr	140
Db	373	AAGTGAAGTTTCATCGCGCTGAACCTTCCCTCCGACGGCCCCCTGATGCAGAAAGAGACC	432
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIlySGlyGlu	160
Db	433	ATGGGCTGGAGAGCCTCCACCGAGCGCTGTACCCCCCGCAGCGCGCTGTGAAGGGCGAG	492
QY	161	IleHisIlyAlaLeuIlySleuIlyAspGlyGlyHisIlyTyrLeuValGluPheIlySerIle	180
Db	493	ACCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTGTGAGTTCAAGTCCATC	552
QY	181	TyrMetAlaIlySlySProValGlnLeuProGlyTyrTyrTyrValAspSerIlySleuAsp	200
Db	553	TACATGGCCCAAGAGCCCGCTGACGTGCCCGGCTACTACTACGTGAGCCCAAGCTGGAC	612
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db	613	ATCACTCCCAACAAGAGACTACCATCTGTGAGACAGTACGAGCGCACCGAGGGCCGC	672
QY	221	HisHisIleuPheLeu	225
Db	673	CACCACCTGTTCTTG	687

```

RESULT 64
US-10-006-922-45
; Sequence 45, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1e1 Chromophores/Fluorophores and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 898
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-45

```

Alignment Scores:	6.18e-139	Length:	898
Pred. No.:	1184.00	Matches:	219
Score:	98.22%	Conservative:	2
Percent Similarity:	97.33%	Mismatches:	4
Best Local Similarity:	97.53%	Indels:	0
Query Match:	5	Gaps:	0
ID:			

US-10-006-922A-12 (1-225) x US-10-006-922-45 (1-898)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20

DB 93 ATGAGGCTTTCAGAATGTTATCAAGAGGTTCAAGGTTTAAAGTTCGCAATGGAAGA 152







```

; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09//866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09//794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 678
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
;   OTHER INFORMATION: "T1"
US-10-931-304-5

```

Alignment Scores:	
Pred. No.:	4.47e-136
Score:	1160.00
Percent Similarity:	97.33%
Best Local Similarity:	96.00%
Query Match:	95.55%
DB:	9
Length:	678
Matches:	216
Conservative:	3
Mismatches:	6
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) x US-10-931-304-5 (1-678)

OY		1	MetArgSerSerLysAsnValIleGlyGluPheMetArgPheLeuValArgMetGluGly	20
Db		1	ATGGCCTCCTCCGAGACGTCAACAAGATTCAATCGCTTCAAGGTCCCATGGAGGC	60
OY		21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db		61	TCCGTGAACGCCCACGAGTTCAAGATCGAGGGCGAGGGCCGCCCTACGAGGC	120
OY		41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db		121	ACCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCTTCCTCGCTGGACATC	180
OY		61	LeuSerProGlnPheGluTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db		181	CTGTCCCCCAGTTCAGTAGCGGTCCAAGTGTAAGTGAAGCACC CGCCGACATCCCC	240
OY		81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db		241	GACTACAAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAAGCGCGTGAATTGAG	300
OY		101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db		301	GACGGCGCGGTGTGACCGCTGACCCAGACTCTCCCTGCAGAGCGGCTCTCATCTAC	360
OY		121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db		361	AAGGTGAAGTTCAATCGGCGTGAACCTTCCCCTCCGACGGCCCCGTAA TGCAAGAAGACT	420
OY		141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db		421	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCCGCGACGGCGGTGTGAAGGGCGAG	480
OY		161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db		481	ATCCACAAGGCCCTGAAGCTGAAGAGCGGGGCCAATACTGCTGAGTTCAAGTTCATC	540
OY		181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db		541	TACATGGCCAGAAGACCCGTGACGTGCCCGGCTACTACTACGTGACTCCAGCTGGAC	600
OY		201	IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg	220
Db		601	ATCACTTCCCAACAAGAGACTAACCATCTGTGAGACAGTACGAGCGCGCCGAGGGCCGC	660
OY		221	HisHisLeuPheLeu	225
Db		661	CACCACTGTTCCTG	675

```

RESULT 67
US-10-844-064A-3
; Sequence 3, Application US/10844064A
; Publication No. US20050149994A1
; GENERAL INFORMATION:
; APPLICANT: Bevis, Brooke
; APPLICANT: Glick, Benjamin
; TITLE OF INVENTION: RAPIDLY MATURING FLUORESCENT PROTEINS AND METHODS FOR USING THE SAME
; FILE REFERENCE: 092234-9006
; CURRENT APPLICATION NUMBER: US/10/844,064A
; CURRENT FILING DATE: 2004-05-11
; PRIOR APPLICATION NUMBER: PCT/US02/40539
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: US 60/341,723
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Discosoma
US-10-844-064A-3

```

Alignment Scores:	
Pred. No.:	4.72e-136
Score:	1160.00
Percent Similarity:	97.33%
Best Local Similarity:	96.00%
Query Match:	95.55%
DB:	9
	Gaps:
US-10-006-922A-12 (1-225) x US-10-844-064A-3 (1-704)	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

QY	1	MeLArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	20	ATGSGCTCTCCGAGACGTCAACAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGC	79
QY	21	ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	80	TCCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCCGCCCTACGAGGGC	139
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle	60
Db	140	ACCAGACCGCCAAAGTGAAGGTGAACAAAGGGCGGCCCTTGCTTGGGACATC	199
QY	61	LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	200	CTGTCCCCCAGTTCAGTACGGCTCAAGTGTACGTGAAGCACCCGCGACATCCCC	259
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu	100
Db	260	GACTACAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATTGACTTCGAG	319
QY	101	AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	320	GACGGCGCGTGTGACCGGTGACCCAGACTCTCTCCGACGAGACGGCTCTCATCTAC	379
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	380	AAAGTGAAGTTCAATCGGCGTGAATCTCCCTCCGACGGCCCCGTAATGCAGAGAAGACT	439
QY	141	MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	440	ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCCGACGCGCGTGTGAAGGGCGAG	499
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	500	ATCCACAAGGCCCTGAAGCTGAAGAGCGGGGCCAATACTGTGTGAGTTCAAGTCCATC	559
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	560	TACATGGCCACAAGACCCCGTGCAGCTGCCCGCTACTACTACGTGACCTCCAAGCTGGAC	619



```

QY 201 ILeThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 620 ATCACCCTCCCAACGACGACTACACCATCGTGCAGCACTACGAGCGCGCGAGGGCCGC 679

QY 221 HisHisLeuPheLeu 225
Db 680 CACCACCCTGTTCCCTG 694

RESULT 68
US-10-423-688A-40
; Sequence 40, Application US/10423688A
; Publication No. US20040078148A1
; GENERAL INFORMATION:
; APPLICANT: Los Alamos National Laboratory
; APPLICANT: Waldo, Geoffrey S.
; TITLE OF INVENTION: Directed Evolution Methods for Improving Polypeptide Folding and
; TITLE OF INVENTION: Solubility and Superfolder Fluorescent Proteins Generated Thereb
; FILE REFERENCE: S-100, 608
; CURRENT APPLICATION NUMBER: US/10/423,688A
; CURRENT FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: 10/132,067
; PRIOR FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
US-10-423-688A-40

Alignment Scores:
Pred. No.: 1.53e-132 Length: 678
Score: 1132.00 Matches: 208
Percent Similarity: 95.56% Conservative: 7
Best Local Similarity: 92.44% Mismatches: 10
Query Match: 93.25% Indels: 0
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-423-688A-40 (1-678)

QY 1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20
Db 1 ATGAGTCTTCCGAGGATGTTATCAAGGAGTTCATGAGGTTTAAAGTTCAATGGAAGCA 60

QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
Db 61 TCGGTCAATGGCGACGAGTTTGAATAGAAGGCCAAGAGGAGGCGCATACGAAGGC 120

QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60
Db 121 ACCCAGAACGTAAAGCTTAAGTTAACTAAGGGGGGACCTTTGCCATTTGCTGGATATT 180

QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80
Db 181 TTGTCAACCAATTTCACTATGGAAGCAAGGTATATGTCAAGCAACCTGCCGACATACCA 240

QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnDheGlu 100
Db 241 GACTATATAAAGCTGTCAATTTCTGAAAGATTATAATGGGAAAGGGTCATGAACTTTGA 300

QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGTGGCGTCGCTACTGTAACCAAGATTCAGTTTGAAGAGATGGCTGTTGATCTAC 360

QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140
Db 361 AAGGTCAAGTTCACTGGCGTGAACCTTCTCCGATGGAACCTGTATGCAAAAGAGACA 420

QY 141 MetGlyTyrGlyAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160
Db 421 ATGGGCTGGGAACCGACGACTGAGCGCTTGTGATCTCTCGTGATGGCGGTGTTGAAGAGAT 480

```

```

Qy      161 ILeHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180
      |||||
Db      481 ATTcATaAGGCTGTGAAGCTGAAGAAGCGGTGGTCATTACTAGTTGATATCAAAAGTATT 540

Qy      181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200
      |||||
Db      541 TAcATGCAAGAAGAGCCCTGTGCAGCTACCAAGGTACTACTAGTTGACTCCAAACTGGAT 600

Qy      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
      |||||
Db      601 ATAACAaACCAcCAAGAcTATAcATcGTTGAGcAGTATGAAGAcCGAcGAGGAcCGc 660

Qy      221 HisHisLeuPheLeu 225
      |||||
Db      661 cAcCATcTGTTCcTT 675

RESULT 69
US-10-121-258-7
; Sequence 7, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "dimer2"
US-10-121-258-7

Alignment Scores:
Pred. No.:      3.77e-131      Length:      681
Score:          1121.00      Matches:      207
Percent Similarity: 95.96%      Conservative: 7
Best Local Similarity: 92.83%      Mismatches: 9
Query Match:    92.34%      Indels:      0
DB:             5      Gaps:      0

US-10-006-922A-12 (1-225) x US-10-121-258-7 (1-681)
Qy      3 SerSerLysAsnValIleLeuGluPheMetArgPheLysValArgMetGluGlyThrVal 22
      |||||
Db      10 TCCTCCGAGAcGTcATCAAGAAGTTCATGCGCTTCAAGGTGGcATGAAGGCTCCGTG 69

Qy      23 AsnGlyHisGluPheGlnIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42
      |||||
Db      70 AACGGCCAcGAGTTCGAcATCGAGGGCGAGGGCGCGCCcCTAcGAGGGcAcCCAcG 129

Qy      43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIleLeuSer 62
      |||||
Db      130 ACCGCCAAcGTGAAGGTGAACCAAGGGCGCGCCcCTGCGCTTGGAcCATCCcGAcTAc 189

Qy      63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82
      |||||
Db      190 CCCAcGTTCCAGTAcGGGTCCAAGCGTAcGTAAGAcAcCCCGCGAcCATCCCGAcTAc 249

Qy      83 LysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGluAspGly 102
      |||||
Db      250 AAGAAGCTGTcCTTCCcCCCGAcGGcCTTCAAGTGGGAAGCGCGGTGAATGACTTCAGAGAcCGc 309

```







Percent Similarity:	95.96%	Conservative:	7
Best Local Similarity:	92.83%	Mismatches:	9
Query Match:	92.34%	Indels:	0
DB:	10	Gaps:	0

US-10-006-922A-12 (1-225) x US-11-052-001-5 (1-1809)

QY	3	SerSerLybAsnValIleIysGluPheMetArgPheIysValArgMetGluIYThrVal	22
Db	424	TCCTCCGAGGACGTCATCAAGAAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCGTG	483
QY	23	AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluIYHisAsn	42
Db	484	AACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCTTACGAGGGCACCCAG	543
QY	43	ThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIleLeuSer	62
Db	544	ACCGCCAAGCTGAAGGTACCAAGGGCGGGCCCCCTGCGCTTGCCTGGGACATCCTGTCC	603
QY	63	ProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIleProAspTyr	82
Db	604	CCCCAGTTCAGTACGGGCTCCAAAGCGGTACGTGAGACACCCCGGACATCCCGGACTAC	663
QY	83	LysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGluAspGly	102
Db	664	AAGAAGCTGCTCTTCCCGAGGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCAGGACGGC	723
QY	103	GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIysVal	122
Db	724	GGCGTGTGACCGTGAACCAAGAACTCTCCCTGCAGAGCGGACGCTGATCTCAAGGTG	783
QY	123	LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThrMetGly	142
Db	784	AAGTTCGGCGGACCAACTTCCCCCGAGCGGCCGTAATGCAAGAAGACCATGGGC	843
QY	143	TrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIleHis	162
Db	844	TGGAGGGCTTCCACCGAGCGCTGTACCCTCCGCGACGGCGGTCTGAAGGGCGAGATCCAC	903
QY	163	LysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIleTyrMet	182
Db	904	CAGGCCCTGAAGCTGAAGGACGGCGGCACTACTGTTGAGTTCAAGACCATCTACATG	963
QY	183	AlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAspIleThr	202
Db	964	GCCAAGAAGCCCGTGAAGTGCCTGGCTACTACTAGTGAACCAAGCTGGACATCACC	1023
QY	203	SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIYArgHisHis	222
Db	1024	TCCCACAACGAGACTACACCATCGTGAACAGTACGAGCGCTCCGAGGGCGCCACAC	1083
QY	223	LeuPheLeu 225	
Db	1084	CTGTTCTCTG 1092	
RESULT 72			
US-10-006-922-43			
; Sequence 43, Application US/10006922			
; Publication No. US20020197676A1			
; GENERAL INFORMATION:			
; APPLICANT: Lukyanov, Sergey A			
; APPLICANT: Fradkov, Arcady F.			
; APPLICANT: Labas, Yulii A.			
; APPLICANT: Matz, Mikhail V.			
; APPLICANT: Tersikh, Alexey			
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and			
; TITLE OF INVENTION: Methods for Using the Same			
; FILE REFERENCE: CLON-035CIP			
; CURRENT APPLICATION NUMBER: US/10/006,922			
; CURRENT FILING DATE: 2001-12-04			
; PRIOR APPLICATION NUMBER: 09/120,330			
; PRIOR FILING DATE: 1998-12-11			
; PRIOR APPLICATION NUMBER: 09/457,898			

```

; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
;   LENGTH: 678
;   TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: hybrid construct
US-10-006-922-43

```

Alignment Scores:	
Pred. No.:	6.7e-131
Score:	1119.00
Percent Similarity:	96.89%
Best Local Similarity:	91.56%
Query Match:	92.17%
DB:	5
Length:	678
Matches:	206
Conservative:	12
Mismatches:	7
Indels:	0
Gaps:	0

US-10-006-922A-12 (1-225) X US-10-006-922-43 (1-678)

QY	1	MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly	20
Db	1	ATGAGCTGCAGCAAGACGTGATCAAGGAGTTCATGCGGTTCAAGGTGCGATGGAGGGC	60
QY	21	ThrValAsnGluYHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly	40
Db	61	ACCGTGAACGCGCACGAGTTCGAGATCAAGGGCGAGGGCGGGCCCTTACGAGGC	120
QY	41	HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle	60
Db	121	CACTGCAGCGTGAAGCTCATGTGTGACCAAGGGCGGCCCTCCCTTCGCTTCGACATC	180
QY	61	LeuSerProGluPheGluTyrGlySerLysValTyrValLysHisProAlaAspIlePro	80
Db	181	CTCAGCCCCCAGTTCCAGTACGGCAGCAAGGTGTACGTGAAGCACCCCGGCACATCCCC	240
QY	81	AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu	100
Db	241	GACTACAAGAAAGCTCAGCTTCCCCAGGGCTTCAAGTGGAGCGGTGATGAATTTCAG	300
QY	101	AspGlyGlyValValThrValThrglnAspSerSerLeuGlnAspGlyCysPheIleTyr	120
Db	301	GACGGCGCGTGTGACCGGTGACGAGACAGCAGCCTCAAGAGCGGTGCTCATCTAC	360
QY	121	LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr	140
Db	361	GAGGTGAAGTTCATCGGCGTGAACCTTCCCCAGCGAGCGCCCGTGATGCAGCGCGGACC	420
QY	141	MetGlyTyrPgluAlaSerThrgluArgLeuTyrProArgAspGlyValLeuLysGlyGlu	160
Db	421	CGGGCGTGGAGGCCAGCAGCAGCGGCTCTACCCCCGGAGCGGCGTGTCTCAAGGGCGAC	480
QY	161	IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle	180
Db	481	ATCCACATGCGCCCTCGGCTCGAGGGCGGGCCGACACTACCTGTGAGTTCAAGACATC	540
QY	181	TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp	200
Db	541	TACATGGCCAGAAGACCCCGTCAGACTCCCCGGCTACTACTACGTGACAGCAAGCTCGAC	600
QY	201	IleThrSerHisAsnGluAspTyrThrIleValGluGluTyrGluArgThrgluArg	220
Db	601	ATCACCAAGCCACAAGAGACTACACCATCTGTGAGCAGTACGAGCGGACCGAGGGCGCGG	660



Oy 221 H1SH1SLeuPheLeu 225  
Db 661 CACCACCTCTCCTC 675

RESULT 73

US-10-081-864-11

/ Sequence 11, Application US/10081864  
/ Publication No. US20030022287A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Lukyanov, Sergey  
/ APPLICANT: Lukyanov, Konstantin  
/ APPLICANT: Yanushevich, Yuriy  
/ APPLICANT: Savietsky, Alexandr  
/ APPLICANT: Pradkov, Arcady  
/ TITLE OF INVENTION: NO. US20030022287A1 Aggregating Fluorescent Proteins and  
/ TITLE OF INVENTION: Methode for Using the Same  
/ FILE REFERENCE: CLON-067  
/ CURRENT APPLICATION NUMBER: US/10/081,864  
/ CURRENT FILING DATE: 2002-06-19  
/ PRIOR APPLICATION NUMBER: 10/006,922  
/ PRIOR FILING DATE: 2001-12-04  
/ PRIOR APPLICATION NUMBER: 60/270,983  
/ PRIOR FILING DATE: 2001-02-21  
/ NUMBER OF SEQ ID NOS: 30  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 11  
/ LENGTH: 678  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: hybrid coding sequence  
US-10-081-864-11

Alignment Scores:  
Pred. No.: 6.7e-131 Length: 678  
Score: 1119.00 Matches: 206  
Percent Similarity: 96.89% Conservative: 12  
Best Local Similarity: 91.56% Mismatches: 7  
Query Match: 92.17% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-081-864-11 (1-678)

Oy 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGAGCTGCAGCAAGAACGTGATCAAGAGTTCAATGCGGTTCAAGGTGCGGATGAGGGC 60  
Oy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
Db 61 ACCGTGAACGGCCACGAGTTCGAGATCAAGGGCGAGGGCGCGCCCTACGAGGCG 120  
Oy 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60  
Db 121 CACTGCAGCGTGAACTCATGGTGAACCAAGGGCGGCCCTTCCTTCGCTTCGACATC 180  
Oy 61 LeuSerProGlnPheGlnTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80  
Db 181 CTCAGCCCCCAGTTCAGTACGGCAAGGTGATGTAAGCACCCCGCCGACATCCCC 240  
Oy 81 AspTyrLySLeuSerPheProGluGlyPheLySTyrGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAAGCTCAGCTTCCCGAGGGCTTCAAGTGGAGCGGGGTGATGAACCTTCAG 300  
Oy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGTGACCGTGAAGCAAGACAGCCTCAAGACGGCTTCATCATC 360  
Oy 121 LySValLySPhelIleGlyValAsnPheProSerAspGlyProValMetGlnLySThr 140  
Db 361 GAGGTGAAGTTCATCGCGGTGAACCTTCCCAAGCGGCCCGGTGATGACGCGCGGACC 420  
Oy 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160

Db 421 CGGGCTGGAGGCCAGCAGCGAGCGGCTTACCCCCGGGAGCGGCGTGTCTCAAGGGCGAC 480  
Oy 161 IleHisLySValLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
Db 481 ATCCACATGCGCTCCGGCTCGAGGGCGGGCCCACTACCTCGTGAAGTCAAGACATC 540  
Oy 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCCAAAGAGCCCGTGACGCTCCCGGCTACTACTACGTGACAGCAAGCTCGAC 600  
Oy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACCAGCCCAAGAGAGACTACACCATCTGAGCAGTACGAGCGGACCGAGGGCCGG 660  
Oy 221 H1SH1SLeuPheLeu 225  
Db 661 CACCACCTCTCCTC 675

RESULT 74

US-10-931-304-82

/ Sequence 82, Application US/10931304  
/ Publication No. US20050196768A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Tsien, Roger  
/ APPLICANT: Campbell, Robert  
/ APPLICANT: Baird, Geoffrey  
/ TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
/ TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
/ FILE REFERENCE: 39754-0831CP2CP3  
/ CURRENT APPLICATION NUMBER: US/10/931,304  
/ CURRENT FILING DATE: 2004-08-30  
/ PRIOR APPLICATION NUMBER: 10/209,208  
/ PRIOR FILING DATE: 2002-07-29  
/ PRIOR APPLICATION NUMBER: 10/121,258  
/ PRIOR FILING DATE: 2002-04-10  
/ PRIOR APPLICATION NUMBER: 09/866,538  
/ PRIOR FILING DATE: 2001-05-24  
/ PRIOR APPLICATION NUMBER: 09/794,308  
/ PRIOR FILING DATE: 2001-02-26  
/ NUMBER OF SEQ ID NOS: 110  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 82  
/ LENGTH: 705  
/ TYPE: DNA  
/ ORGANISM: Artificial Sequence  
/ FEATURE:  
/ OTHER INFORMATION: Polynuc. of Polypeptide variant "dimer2.2MM(dimer3) (dtomato)"  
US-10-931-304-82

Alignment Scores:  
Pred. No.: 7.57e-128 Length: 705  
Score: 1095.00 Matches: 201  
Percent Similarity: 94.57% Conservative: 8  
Best Local Similarity: 90.95% Mismatches: 12  
Query Match: 90.20% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-931-304-82 (1-705)

Oy 5 LySAsnValIleLySGluPheMetArgPheLyValArgMetGluGlyThrValAsnGly 24  
Db 16 GAGGAGGTCAATCAAGAGTTCATGCGCTTCAAGGTGCGGATGAGGGCTCATGAACGGC 75  
Oy 25 HisGluPheGluIleGluGlyGluGlyGlyArgProTyrGluGlyHisAsnThrVal 44  
Db 76 CACGAGTTCAAGATCGAGGGCGAGGGCGGCGCCCTTACGAGGGCAACCAAGACCGCC 135  
Oy 45 LySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGln 64  
Db 136 AAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTGCGGACATCTGTCCCCCAG 195  
Oy 65 PheGlnTyrGlySerLySValTyrValLyHisProAlaAspIleProAspTyrLyS 84



Db 196 TTCATGTACGGCTCCAAGGCGGTACGTGAAGCACCCCGACATCCCCGATTACAAGAG 255  
QY 85 LeuSerPheProGluGlyPheIleSTrpGluArgValMetAsnPhelGluAspGlyGlyVal 104  
Db 256 CTGTCTTCCCGAGGGCTTCAAGTGGAGCGGTGATGAACCTTCAGAGACGGCGGTCTG 315  
QY 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPhe 124  
Db 316 GTGACCGGTGACCAAGGACTCTCTCCCTGCAGAGACGGCAGCTGATCTACAGGTGAAGATG 375  
QY 125 IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu 144  
Db 376 CGCGGCACCAACTTCCCCCCCCGACGGCCCCGTATGCAAGAAAGACCATGGGCTGGAG 435  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyIleHisLysAla 164  
Db 436 GCCTCCACCGAGCGCTGTACCCCCCGCAGCGGTGCTGAAGGGCGAGATCCACGAGCC 495  
QY 165 LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLys 184  
Db 496 CTGAAGCTGAAGACGGCGGCGCACTACTGCTGAGGTCAAGACCATCTAATGCGCAAG 555  
QY 185 LysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHis 204  
Db 556 AAGCCCGTGCACTGCCCGGCTACTACTACGTGAGACACCAAGCTGACATCACCTCCAC 615  
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHisLeuPhe 224  
Db 616 AACGAGACTACCATCTGTGAAACAGTACGAGCGCTCCGAGGGGCCCAACCACTGTTC 675  
QY 225 Leu 225  
Db 676 CTG 678

RESULT 75  
US-10-931-304-107  
; Sequence 107, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 107  
; LENGTH: 1431  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "cdTomato"  
US-10-931-304-107

Alignment Scores:  
Pred. No.: 2.03e-127 Length: 1431  
Score: 1095.00 Matches: 201  
Percent Similarity: 94.57% Conservative: 8  
Best Local Similarity: 90.95% Mismatches: 12  
Query Match: 90.20% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-107 (1-1431)  
QY 5 LysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGly 24  
Db 16 GAGAGGTTCATCAAGAAGTTCAATCGCTTCAAGGTGCGCATGGAGGGCTTCATGAACGGC 75  
QY 25 HisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrVal 44  
Db 76 CACGAGTTGAGATCGAGGGCGAGGGCGAGGGCGCCCTACGAGGGGCAACCAAGACCGCC 135  
QY 45 LysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIleLeuSerProGln 64  
Db 136 AAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTTCGCTGGACATCTTCCCCCAG 195  
QY 65 PheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLys 84  
Db 196 TTCATGTACGGCTCCAAGGCGGTACGTGAAGCACCCCGCAGACATCCCCGATTACAAGAG 255  
QY 85 LeuSerPheProGluGlyPheIleSTrpGluArgValMetAsnPhelGluAspGlyGlyVal 104  
Db 256 CTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCAGAGACGGCGGTCTG 315  
QY 105 ValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPhe 124  
Db 316 GTGACCGGTGACCAAGGACTCTCTCCCTGCAGAGACGGCAGCTGATCAAGGTGAAGATG 375  
QY 125 IleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGlu 144  
Db 376 CGCGGCACCAACTTCCCCCCCCGACGGCCCCGTATGCAAGAAAGACCATGGGCTGGAG 435  
QY 145 AlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyIleHisLysAla 164  
Db 436 GCCTCCACCGAGCGCTGTACCCCCCGCAGCGGTGCTGAAGGGCGAGATCCACGAGCC 495  
QY 165 LeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLys 184  
Db 496 CTGAAGCTGAAGACGGCGGCGCACTACTGCTGAGTTCAAGACCATCTAATGCGCAAG 555  
QY 185 LysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHis 204  
Db 556 AAGCCCGTGCACTGCCCGGCTACTACTGAGACACCAAGCTGACATCACCTCCAC 615  
QY 205 AsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHisLeuPhe 224  
Db 616 AACGAGACTACCATCTGTGAAACAGTACGAGCGCTCCGAGGGGCCCAACCACTGTTC 675  
QY 225 Leu 225  
Db 676 CTG 678

RESULT 76  
US-10-006-922-17  
; Sequence 17, Application US/10006922  
; Publication No. US20020197676A1  
; GENERAL INFORMATION:  
; APPLICANT: Lukyanov, Sergey A  
; APPLICANT: Fradkov, Arcady F.  
; APPLICANT: Labas, Yulii A.  
; APPLICANT: Matz, Mikhail V.  
; APPLICANT: Teresikh, Alexey  
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and  
; FILE REFERENCE: CLON-035CIP  
; CURRENT APPLICATION NUMBER: US/10/006,922  
; CURRENT FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: 09/120,330  
; PRIOR FILING DATE: 1998-12-11  
; PRIOR APPLICATION NUMBER: 09/457,898  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,144  
; PRIOR FILING DATE: 1999-12-09  
; PRIOR APPLICATION NUMBER: 09/458,477  
; PRIOR FILING DATE: 1999-12-09



```

; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-17
```

```

Alignment Scores:
Pred. No.:      1.62e-126      Length:      876
Score:          1085.50        Matches:     199
Percent Similarity: 95.58%      Conservative: 17
Best Local Similarity: 88.05%    Mismatches:   9
Query Match:    89.42%         Indels:      1
DB:             5              Gaps:          1
```

US-10-006-922a-12 (1-225) x US-10-006-922-17 (1-876)

```

QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
      |||
Db      45 ATGAGTTGTTCCAAGAATGATCAAGAGTTTCATGAGTTCAAGGTTCTGATGAGA 104

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||
Db      105 ACGGTCAATGGGCAAGAGTTGAATAAAGCGAAGGTGAAGGAGCCTTACGAAGT 164

QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60
      |||
Db      165 CACTGTTCCGTAAGCTTATGGTAAACCAAGGGTGAGCCTTTGCCATTGCTTTGATATT 224

QY      61 LeuSerProGluPheGluTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
      |||
Db      225 TTGTCAACCAATTTCAATGATGAGCAAGGATATATCAAAACCCCTGCCGACATACCA 284

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||
Db      285 GACTATAAAAAAGCTGTCATTTCTGACGAGATTAAATGGGAAAGGTCATGAACTTTGAA 344

QY      101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120
      |||
Db      345 GACGGTGGCGGTGTTACTGTATCCCAAGATTCACAGTTTGAAAGACGGCTGTTTCATCTAC 404

QY      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGluNlySlyThr 140
      |||
Db      405 GAGGTCAAGTTCAATGGGGTGAACCTTCTTCTGATGACCTGTTATGCAAGAGAGACA 464

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
      |||
Db      465 CGGGCTGGGAAGCCAGCTCTGAGCGTTTGTATCCTCGTATGGGGTGCTGAAAGGAGAC 524

QY      161 IleHisLySAlaLeuLySLeuLySAspGlyGlyHisTyrLeuValGluPheLySserIle 180
      |||
Db      525 ATCCATATGGCTCTGAGGCTGGAAGAGCGGCCATTACCTCGTTGAATTCAAAAGTATT 584

QY      181 TyrMetAlaLySLeuSPro--ValGluLeuProGlyTyrTyrTyrValAspSerLySLeu 199
      |||
Db      585 TACATGGTAAGAAGCCTTCAGTGCAGTGGCCAAGCTACTATTAAGTTGACTCCAAACTG 644

QY      200 AspIleThrSerHisAsnGluAspTyrThrIleValGluGluTyrGluArgThrGluGly 219
      |||
Db      645 GATATGACGAGCCACACAGAAATTACACAGTCGTTGAGCAGATGAAAAAACCAAGGGA 704

QY      220 ArgHisHisLeuPheLeu 225
      |||
Db      705 CGCCACCATCCGTTCAAT 722

RESULT 77
US-10-161-403-39
; Sequence 39, Application US/10161403
```

```

; Publication No. US20030119104A1
; GENERAL INFORMATION:
; APPLICANT: Perkins, Edward
; APPLICANT: Perez, Carl
; APPLICANT: Lindenbaum, Michael
; APPLICANT: Greene, Amy
; APPLICANT: Leung, Josephine
; APPLICANT: Fleming, Elena
; APPLICANT: Stewart, Sandra
; APPLICANT: Shellard, Joan
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS
; FILE REFERENCE: 24601-420
; CURRENT APPLICATION NUMBER: US/10/161,403
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/294,758
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 60/366,891
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Discosoma species
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (45)...(737)
; OTHER INFORMATION: Nucleotide sequence encoding red fluorescent
; OTHER INFORMATION: protein (FP593)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF272711
; DATABASE ENTRY DATE: 2000-09-26
US-10-161-403-39
```

```

Alignment Scores:
Pred. No.:      1.62e-126      Length:      876
Score:          1085.50        Matches:     199
Percent Similarity: 95.58%      Conservative: 17
Best Local Similarity: 88.05%    Mismatches:   9
Query Match:    89.42%         Indels:      1
DB:             6              Gaps:          1
```

US-10-006-922a-12 (1-225) x US-10-161-403-39 (1-876)

```

QY      1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20
      |||
Db      45 ATGAGTTGTTCCAAGAATGATCAAGAGTTTCATGAGTTCAAGGTTCTGATGAGA 104

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
      |||
Db      105 ACGGTCAATGGGCAAGAGTTGAATAAAGCGAAGGTGAAGGAGCCTTACGAAGT 164

QY      41 HisAsnThrValLySLeuLySValThrLySGlyGlyProLeuProPheAlaTyrAspIle 60
      |||
Db      165 CACTGTTCCGTAAGCTTATGGTAAACCAAGGGTGAGCCTTGGCATTGCTTTGATATT 224

QY      61 LeuSerProGluPheGluTyrGlySerLySValTyrValLyHisProAlaAspIlePro 80
      |||
Db      225 TTGTCAACCAATTTCAATGATGAGCAAGGATATATGTCAAACCCCTGCCGACATACCA 284

QY      81 AspTyrLySLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100
      |||
Db      285 GACTATAAAAAAGCTGTCATTTCTGAGGAGATTAAATGGGAAAGGTCATGAACCTTTGAA 344

QY      101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120
      |||
Db      345 GACGGTGGCGGTGTTACTGTATCCCAAGATTCCAGTTTGAAAGACGGCTGTTTCATCTAC 404

QY      121 LysValLySPheIleGlyValAsnPheProSerAspGlyProValMetGluNlySlyThr 140
      |||
Db      405 GAGGTCAAGTTCAATGGGGTGAACCTTCTTCTGATGACCTGTTATGACAGAGAGACA 464

QY      141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160
```



Db 465 CGGGGCTGGGAAGCCAGCTCTGAGCGTTGTATCTCGTGATGGGGTCTGAAAGGAGAC 524  
Qy 161 ILEHISLYSALALEULYSLEULYSASPGLYGlyHISTYrLEuValGluPheLYSserILE 180  
Db 525 ATCCATATGGCTCTGAGGCTGGAAGGAGGGCCATTACCTCGTTGAATTCAAAAGTATT 584  
Qy 181 TYrMetAlAlYSLSPro--ValGlnLeuProGlyTYrTYrTYrValAlAsPserLYSLeu 199  
Db 585 TACATGTAAAGAACCTTCAGTGCAGTGCAGGCTACTATTATGTGACTCCAAACTG 644  
Qy 200 ASPLeThrSerHISAsnGluAspTYrThrILEValGluGlnTYrGluArgThrGluGly 219  
Db 645 GATATGACGAGCCACAACGAAGATTACAGTCGTTGAGCAGTATGAATAAACCCAGGGA 704  
Qy 220 ARGHISHILEUPheLeu 225  
Db 705 CGCCACCATCCGTTCAATT 722

RESULT 78

US-11-006-076-39  
; Sequence 39, Application US/11006076  
; Publication No. US20050181506A1  
; GENERAL INFORMATION:  
; APPLICANT: Perkins, Edward  
; APPLICANT: Perez, Carl  
; APPLICANT: Lindenbaum, Michael  
; APPLICANT: Greene, Amy  
; APPLICANT: Leung, Josephine  
; APPLICANT: Fleming, Elena  
; APPLICANT: Stewart, Sandra  
; APPLICANT: Shellard, Joan  
; TITLE OF INVENTION: CHROMOSOME-BASED PLATFORMS  
; FILE REFERENCE: 17084-022002/420B  
; CURRENT APPLICATION NUMBER: US/11/006,076  
; PRIOR FILING DATE: 2004-12-06  
; PRIOR APPLICATION NUMBER: 60/294,758  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/366,891  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Drososoma species  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (45)..(737)  
; OTHER INFORMATION: Nucleotide sequence encoding red fluorescent  
; OTHER INFORMATION: protein (FP593)  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank AF272711  
; DATABASE ENTRY DATE: 2000-09-26  
US-11-006-076-39

Alignment Scores:  
Pred. No.: 1.62e-126 Length: 876  
Score: 1085.50 Matches: 199  
Percent Similarity: 95.58% Conservative: 17  
Best Local Similarity: 88.05% Mismatches: 9  
Query Match: 89.42% Indels: 1  
DB: 10 Gaps: 1

US-10-006-922A-12 (1-225) x US-11-006-076-39 (1-876)

Qy 1 MetArgSerSerLYSAsnValILELYSGluPheMetArgPheLYSValArgMetGluGly 20  
Db 45 ATGAGTTGTTCCAAGAAATGTGATCAAGGAGTTCAAGGTTCAAGGTTCTATGGAAGGA 104  
Qy 21 ThrValAsnGlyHISGluPheGluILEGluGlyGluGlyGluGlyArgProTYrGluGly 40  
Db 105 ACGGTCAATGGGCACGAGTTTGAATAAAGCGAAGGTGAAGGAGGCGCTTACGAAGGT 164

Qy 41 HISAsnThrValLYSLeuLYSValThrLYSGlyGlyProLeuProPheAlaTrpAspILE 60  
Db 165 CACTGTTCCTTAAGCTTATGTGTAACCAAGGGTGACCTTTGCCATTGCTTTGATATT 224  
Qy 61 LeuSerProGlnPheGlnTYrGlySerLYSValTYrValLYSHISProAlaAspILEPro 80  
Db 225 TTGTACCAACAATTTCAGTATGGAAGCAAGGTATATGTCAACACCCTGCCACATACCA 284  
Qy 81 ASPTYrLYSLSLeuSerPheProGluGlyPheLYSTrPGluArgValMetAsnPheGlu 100  
Db 285 GACTATATAAAGCTGTCAATTCTGAGGATTAAATGGGAAGGTCATGAACCTTGA 344  
Qy 101 ASPGlyGlyValThrValThrGlnAsPserSerLeuGlnAspGlyCysPheILETYr 120  
Db 345 GACGTGCGGTGGTTACTGTATCCCAAGATTCCAGTTTGAAAGACGGCTGTTCACTAC 404  
Qy 121 LYSValLYSPheILEGlyValAsnPheProSerAspGlyProValMetGlnLYSLeuThr 140  
Db 405 GAGGTCAAGTTCAATTGGGGTGAACTTCTCTGATGACCTGTTATGCAAGAGAGACA 464  
Qy 141 MetGlyTYrPGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLYSGlyGlu 160  
Db 465 CGGGGCTGGGAAGCCAGCTCTGAGCGTTGTATCTCGTATGGGGTCTGAAAGGAGAC 524  
Qy 161 ILEHISLYSALALEULYSLEULYSASPGLYGlyHISTYrLEuValGluPheLYSserILE 180  
Db 525 ATCCATATGGCTCTGAGGCTGGAAGGCGGCCATTACCTCGTTGAATTCAAAAGTATT 584  
Qy 181 TYrMetAlAlYSLSPro--ValGlnLeuProGlyTYrTYrTYrValAlAsPserLYSLeu 199  
Db 585 TACATGTAAAGAACCTTCAGTGCAGTGCAGGCTACTATTATGTGACTCCAAACTG 644  
Qy 200 ASPLeThrSerHISAsnGluAspTYrThrILEValGluGlnTYrGluArgThrGluGly 219  
Db 645 GATATGACGAGCCACAACGAAGATTACAGTCGTTGAGCAGTATGAATAAACCCAGGGA 704  
Qy 220 ARGHISHILEUPheLeu 225  
Db 705 CGCCACCATCCGTTCAATT 722

RESULT 79

US-10-724-178-15  
; Sequence 15, Application US/10724178  
; Publication No. US20040137528A1  
; GENERAL INFORMATION:  
; APPLICANT: Odyssey Thera, Inc.  
; APPLICANT: Michnick, Stephen  
; APPLICANT: MacDonald, Marne  
; APPLICANT: Lamerdin, Jane  
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT  
; TITLE OF INVENTION: COMPLEMENTATION ASSAYS  
; FILE REFERENCE: ODDY007  
; CURRENT APPLICATION NUMBER: US/10/724,178  
; PRIOR FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: US 60/461,133  
; PRIOR FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: "mRFP1", an engineered monomeric form of DsRed FP  
; NAME/KEY: CDS  
; LOCATION: (1)..(675)  
; PUBLICATION INFORMATION:  
; AUTHORS: Campbell, R.E. et al  
; TITLE: A monomeric red fluorescent protein  
; JOURNAL: Proc. Nat'l. Acad. Sci.  
; VOLUME: 99



ISSUE: 12  
PAGES: 7877-82  
DATE: 2002-06-11  
RELEVANT RESIDUES: (1)..(675)  
US-10-724-178-15

Alignment Scores:  
Pred. No.: 1.56e-118 Length: 675  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-724-178-15 (1-675)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCCTCCTCCGAGACGTCAAGAGATTCAATCGCTTCAAGGTGCGCATGGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTGCGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAGTTCCAGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGCATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnGlu 100  
DB 241 GACTACTTGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGCGGTGTGACCGTGAACCAAGACTCTCTCTGCAAGACGGCGAGTTCACTTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCACCACTTCCCTCCGACGGCCCGTAAATGCAAGAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCCACCGACGGAGTATACCCCGAGGCGGCCCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCCTACTACGACCGGAGGTCAAGACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCGCTTACAAAGACCGACATCAAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGAGAGACTACCATCGTGAACAGTACGAGCGCGCGAGGGCGCG 660  
QY 221 His 221  
DB 661 CAC 663

RESULT 80  
US-10-121-258-9  
Sequence 9, Application US/10121258  
Publication NO. US20030059835A1  
GENERAL INFORMATION:  
APPLICANT: Telen, Roger  
APPLICANT: Campbell, Robert  
TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME

FILE REFERENCE: UC083.1CP2CP1  
CURRENT APPLICATION NUMBER: US/10/121,258  
CURRENT FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: 09/794,308  
PRIOR FILING DATE: 2001-02-26  
PRIOR APPLICATION NUMBER: 09/866,538  
PRIOR FILING DATE: 2001-05-24  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for windows Version 4.0  
SEQ ID NO 9  
LENGTH: 678  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
OTHER INFORMATION: "mrp1"  
US-10-121-258-9

Alignment Scores:  
Pred. No.: 1.57e-118 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 5 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-121-258-9 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCCTCCTCCGAGACGTCAAGAGATTCAATCGCTTCAAGGTGCGCATGGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGGCGGCCCTTGCCCTGCGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAGTTCCAGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGCATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnGlu 100  
DB 241 GACTACTTGAAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCCAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGCGGTGTGACCGTGAACCAAGACTCTCTCTGCAAGACGGCGAGTTCACTTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCACCACTTCCCTCCGACGGCCCGTAAATGCAAGAGACC 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCCACCGACGGAGTATACCCCGAGGACGGCGCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCCTACTACGACCGGAGGTCAAGACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCGCGCTTACAAAGACCGACATCAAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGAGAGACTACCATCGTGAACAGTACGAGCGCGCGAGGGCGCG 660  
QY 221 His 221



Db 661 CAC 663

RESULT 81

US-10-931-304-9  
; Sequence 9, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "mRFP1"  
US-10-931-304-9

Alignment Scores:  
Pred. No.: 1.57e-118 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-9 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGGCCTCCTCCGAGACGTCAAGGATTCATGCGCTTCAAGGTGCGCATGGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGCGCGCCCTACGAGGCG 120  
QY 41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGCGGCCCCCTGCTTCCCTGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLySHisProAlaAspIlePro 80  
Db 181 CTGTCCCTCAGTTCCAGTACGAGGCTCCAAAGGCTTACGTGAAGCACCCCGACATCCCC 240  
QY 81 AspTyrLyLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACTTGAAAGTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGACAGACGGCGAGTTCATCTAC 360  
QY 121 LySValLyPheIleGlyValAsnPheProSerAspGlyProValMetGlnLySLeThr 140  
Db 361 AAGGTGAAGCTGCGCGGCAACCAACTTCCCTCCGACGGCCCCGTAAATGAGAAGACC 420

QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
Db 421 ATGGGCTGGAGGCTCCACCGAGCGATGTACCCCGAGACGGCGCCCTGAAGGGCGAG 480  
QY 161 IleHisLyAlaLeuLySLeuLyAspGlyGlyHisTyrLeuValGluPheLySerIle 180  
Db 481 ATCAAGATGAGGCTGAAGCTGAAGACGCGCGCCACTACGACGCCGAGCTCAAGACCACC 540  
QY 181 TyrMetAlaLySLeuProValGlnLeuProGlyTyrTyrTyrValAspSerLySLeuAsp 200  
Db 541 TACATGGCCAGAAGACCCCGTGAGCTGCCCGCGCCCTACAAAGACCATCAAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACCCTCCCAACGAGGACTACCATCTGTGAACAGTACGACGCCCGGAGGGCGCG 660  
QY 221 His 221  
Db 661 CAC 663

RESULT 82

US-10-931-304-109  
; Sequence 109, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 109  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mGrapel"  
US-10-931-304-109

Alignment Scores:  
Pred. No.: 1.2e-117 Length: 678  
Score: 1014.00 Matches: 191  
Percent Similarity: 89.14% Conservative: 6  
Best Local Similarity: 86.43% Mismatches: 24  
Query Match: 83.53% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-109 (1-678)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
Db 1 ATGGCCTCCTCCGAGACGTCAAGGATTCATGCGCTTCAAGGTGCGCATGGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40  
Db 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGCGCGCCCTACGAGGCG 120  
QY 41 HisAsnThrValLySLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 ACCCAGACCGCCCAAGCTGAAGGTGACCAAGGCGGCCCCCTGCTTCCCTGGACATC 180



Qy 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAGTTCATGTACGGCTCCAGGCTTACGTGAAGCACCCCGCATCTCCC 240  
Qy 81 AspTyrIysValLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACCTGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCCGGCTGTGACCGTGACCCAGACTCTCCCTGACAGACGGGAGTTTATCTAC 360  
Qy 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCACTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC 420  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCTCCGAGCGGCTGTACCCCGAGACGGCGCGCTGAAGGGCGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCACTACGACCGGAGCCCAAGACC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGCTGCAAGCTGCCCGCGCTTACAAGCTCGACTACAGCTGGAC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGAGACTACACCATCTGTGAACAGTACGAGCGCGCCGAGGGCCGC 660  
Qy 221 His 221  
DB 661 CAC 663

RESULT 83  
US-10-931-304-80  
; Sequence 80, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 80  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "mrFp1.1"  
US-10-931-304-80

Alignment Scores:  
Pred. No.: 2.15e-117 Length: 678  
Score: 1012.00 Matches: 190  
Percent Similarity: 88.69% Conservative: 6  
Best Local Similarity: 85.97% Mismatches: 25

Query Match: 83.36% Indels: 0  
DB: 9 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-931-304-80 (1-678)

Qy 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCTCTCTCCGAGAGCTATCAAGAGATTTCATGCGCTTCAAGGTGCCATGAGGGGC 60  
Qy 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCCCTTACGAGGGC 120  
Qy 41 HisAsnThrValLysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIle 60  
DB 121 ACCGAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCGCTTGGAGACATC 180  
Qy 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAGTTCATGTACGGCTCCAGGCTTCAAGGCTTCAAGTGGAGCGCGTGAATCTGAG 240  
Qy 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACTTGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGTGAATCTGAG 300  
Qy 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGGCTGTGACCGTGACCCAGACTCTCTCCGAGACGGCGGAGTTTATCTAC 360  
Qy 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysIysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCACTTCCCTCCGACGGCCCCGTATGCAAGAAAGACC 420  
Qy 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCTCCGAGCGGATGTACCCCGAGACGGCGCGCTGAAGGGCGAG 480  
Qy 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGGCGGCACTACGACCGGAGGTCAAGACCACC 540  
Qy 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAAGAGCCCGTGAAGCTGCCCGCGCTTACAAGACCGACATCAAGCTGAGC 600  
Qy 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAGAGACTACACCATCTGTGAACAGTACGAGCGCGCGAGGGCCGC 660  
Qy 221 His 221  
DB 661 CAC 663

RESULT 84  
US-10-931-304-105  
; Sequence 105, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308



; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mHoneydew"
US-10-931-304-105

Alignment Scores:
Pred. No.: 9.2e-117 Length: 678
Score: 1007.00 Matches: 188
Percent Similarity: 88.69% Conservative: 8
Best Local Similarity: 85.07% Mismatches: 25
Query Match: 82.95% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-105 (1-678)

QY 1 MetArgSerSerLybAsnValIlelyeGluPhemeTArgPhelysValArgMetGluGly 20
Db 1 ATGGCCCTCCTCCGAGGACGTCATCAAGAGTTCAATGCGCTTCAAGGTCCGATGAGGGC 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCCGCCCTACGAGGGC 120
QY 41 HisAsnThrVallybLeuLybValThrLybGlyGlyProLeuProPheAlaTPrAspIle 60
Db 121 ACCCAGACCGCCAGCTGAAGGTGACCAAGGCGGCCCTGCTGCTTCCCTGGACATC 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
Db 181 CTGTCCCTCAGTTTCATGTGGGCTCCAGGCTTACGTGAAGCACCCGCCGACATCCCC 240
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTPrGluArgValMetAsnPheGlu 100
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGCTGAACTTCGAG 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGCGCGGTGTGACCGTGACCCAGACTCTCTCCCTGCAGAGCGCGGAGTTTATCTAC 360
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140
Db 361 AAGGTGAAGCTGCGCGGACCAACTTCCCTCCGACGGCCCCGTAAATGCAGAAAGAACCC 420
QY 141 MetGlyTPrGluAlaSerThrGluArgLeuTyrProArgAspGlyValleuLybGlyGlu 160
Db 421 ATGGGCTGGGCGGCCACCAACGAGGATGTAACCCGAGAGCGGCGCTTGAAGGGCGAG 480
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180
Db 481 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCCGAGTCAAGACCACC 540
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200
Db 541 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCCCTACAAAGATTGAAGGGAAGCTGAGC 600
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
Db 601 ATCACCCTCCCAACGAGGACTACACCATTCGTGAACAGTACGAGCGCGCGAGGGCGGC 660
QY 221 His 221
Db 661 CAC 663

RESULT 85
US-10-931-304-101
; Sequence 101, Application US/10931304
; Publication No. US20050196768A1

; GENERAL INFORMATION:
; APPLICANT: Telen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mTangerine"
US-10-931-304-101

Alignment Scores:
Pred. No.: 3.94e-116 Length: 678
Score: 1002.00 Matches: 189
Percent Similarity: 88.24% Conservative: 6
Best Local Similarity: 85.52% Mismatches: 26
Query Match: 82.54% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-101 (1-678)

QY 1 MetArgSerSerLybAsnValIlelybGluPhemeTArgPhelysValArgMetGluGly 20
Db 1 ATGGCCCTCCTCCGAGGACGTCATCAAGAGTTCAATGCGCTTCAAGGTCCGATGAGGGC 60
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40
Db 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCCGCCCTACGAGGGC 120
QY 41 HisAsnThrVallybLeuLybValThrLybGlyGlyProLeuProPheAlaTPrAspIle 60
Db 121 ACCCAGACCGCCAGCTGAAGGTGACCAAGGCGGCCCTGCTGCTTCCCTGGACATC 180
QY 61 LeuSerProGlnPheGlnTyrGlySerLybValTyrValLybHisProAlaAspIlePro 80
Db 181 CTGTCCCTCAGTTTCTGTACGGCTCCAAAGGCTTACGTGAAGCACCCCGACATCCCC 240
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLybTPrGluArgValMetAsnPheGlu 100
Db 241 GACTACTTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 300
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
Db 301 GACGGCGCGGTGTGACCGTGACCAACTTCCCTCCGACGGCCCCGTAAATGCAGAAAGAACCC 420
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLybThr 140
Db 361 AAGGTGAAGCTGCGCGGACCAACTTCCCTCCGACGGCCCCGTAAATGCAGAAAGAACCC 420
QY 141 MetGlyTPrGluAlaSerThrGluArgLeuTyrProArgAspGlyValleuLybGlyGlu 160
Db 421 ATGGGCTGGAGGCTCTCTCCGAGCGGATGTACCCCGAGAGCGGCGCTTGAAGGGCGAG 480
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180
Db 481 ATCAAGATGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCCGAGGTCAAGACCACC 540



```

Qy 181 IYMEtAlAlYsYsPVoAlGlnLeuProGlyTYrTYrTYrValAspSerLYsLeuAsp 200
    |||||
Db 541 TACATGGCCAAAGAGCCCGTGACGTGCCGCGCCTTACAAGACCGACATCAAGCTGGAC 600

Qy 201 ILeThrSerHisAsnGluAspTYrThrIleValGlnTYrGluArgThrGluGlyArg 220
    |||||
Db 601 ATCACTCCCAACAAGAGGAACTACACCATCTGGAAATTGTACGAGCGCGCGAGGCGCGC 660

Qy 221 HIs 221
    |||
Db 661 CAC 663

RESULT 86
US-10-931-304-86
; Sequence 86, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsielen, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baltd, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding Daked polypeptide variant "OrS4-9"
US-10-931-304-86

Alignment Scores:
Pred. No.: 4.3e-115 Length: 711
Score: 994.00 Matches: 186
Percent Similarity: 88.84% ConservatIve: 5
Best Local Similarity: 86.51% Mismatches: 24
Query Match: 81.88% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-86 (1-711)

Qy 7 ValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHisGlu 26
    ::|||
Db 34 ATCATCAAGGAGTTCATGCGCTTCAAGGTACGATGAGGGCTCCGTGAACGCCACGAG 93

Qy 27 PheGluIleGluGlyGluGlyGluGlyArgProTYrGluGlyHisValnThrValLysLeu 46
    |||||
Db 94 TTCGAGATTGAGGGCGAGGGCGAGGGCCGCCCTTACGAGGGGACCCAGACGCCAAGTTG 153

Qy 47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
    |||||
Db 154 AAGGTACCAAGGGGTGGCCCCCTGCCCTTGGCTGGACATCTGTCCCCCTCAGTTCAACC 213

Qy 67 TYrGlySerLysValTYrValLysHisProAlaAspIleProAspTYrLysLeuSer 86
    |||||
Db 214 TACGGCTCCAAAGGCTTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCC 273

Qy 87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106
    |||||
Db 274 TTCCCGGAGGCTTCAAGTGGGAACGCGTGATGAACCTTCAGAGGACGGCGGCGTGGTACC 333

```

```

QY      107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIlysValIysPheIleGly 126
      |||
DB      334 GTGACACAGAGACTCTCCCTCGACAGACGGCGAGTTCATCTACAAAGTGAAGCTGCCGGC 393

QY      127 ValAsnProSerAspGlyProValMetGlnIlybIysThrMetGlyTrpGluAlaSer 146
      |||
DB      394 ACCAACTTCCCTCCGACGAGCCCGTAATGACAGAAAGACCATGGGCTGGGAGGCTCC 453

QY      147 ThrGluArgLeuTyrProArgAspGlyValLeuIysGlyIleHisIlybAlaLeuIys 166
      ::|||
DB      454 TCCGAGCGGATGTAACCCCGAGGACGGCGCCCTGAAAGGCGAGATCAAGATGAGGCTGAAG 513

QY      167 LeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIleTyrMetAlaIysIysPro 186
      |||
DB      514 CTGAAGACGCGCGGCACTACGACGCTGAGGTCAAGACCACCTACAAAGGCCAAGAACCC 573

QY      187 ValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAspIleThrSerHisAsnGlu 206
      |||
DB      574 GTGACAGTGCCTCCGCGCTACAAAGTGCACATCAAGCTGACATCACCCTCCCAACGAG 633

QY      207 AspTyrThrIleValGluGlnTyrGluArgThrGluArgHis 221
      |||
DB      634 GACTACACCATCGTGAACAGTACGAACGCGCGGAGGCGCCGAC 678

RESULT 87
US-10-931-304-90
; Sequence 90, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polymuc. encoding polypeptide variant "mFRFP (F2Q6) (mGrape2)"
US-10-931-304-90

Alignment Scores:
Pred. No.: 7.69e-115 Length: 711
Score: 992.00 Matches: 185
Percent Similarity: 88.84% Conservative: 6
Best Local Similarity: 86.05% Mismatches: 24
Query Match: 81.71% Indels: 0
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-90 (1-711)

QY      7 ValIleIysGluPheMetArgPheIysValArgMetGluGlyThrValAsnGlyHisGlu 26
      |||
DB      34 GTCATCAAGAGATTGATGCGCTTCAAGGTGCGCATGAGGGCTCGTGAACGGCCACGAG 93

QY      27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrValIysLeu 46
      |||
DB      94 TTCGAGATCGAGGGCAAGGGCGAGGGCCGCCCTACGAGGGGCAACCCAGCCGCAAGCTG 153

```



QY 47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
DB 154 AAGGTGACCAAGGGCGGCCCTGCTCCCTGGACATCTGTCCCTCAGTTCATG 213  
QY 67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer 86  
DB 214 TACGAGCTCCAAAGCCTTACGTGAAGCACCCCTGACATCCCCGAGTACATGAAGCTGTCC 273  
QY 87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106  
DB 274 TTCCCCGAGGGCTTCAAGTGGAGCGCGGTGTAATTGAGGACGCGCGCGGTGACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGly 126  
DB 334 GTGACCCAGAGACTCTCCCTGACAGACGCGCGAGTTCATCTACAAGGTGAAGCTGCACGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGluAlaSer 146  
DB 394 ACCAACTTCCCTCCGACGGCCCCGTAAATGCAAGAACCATGGGCTGGGAGGCTCC 453  
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166  
DB 454 TCCGAGCGGTTGTACCCCGAGACGCGCGCTGAAGGCGAGGTCAAGATGAGGCTGAAG 513  
QY 167 LeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
DB 514 CTGAAGGACGGGGCCCACTACGACGCCGAGGTCAAGACCACCTACATGGCCCAAGAGCCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
DB 574 GTGACAGTGGCCCGGCCCTTACAAGCTCGACTACAAAGCTGGACATCCTCCCAACAAGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
DB 634 GACTACACCATCTGTGAACAGTACGAGCGCGCCGAGGGCGCCAC 678

RESULT 88  
US-10-931-304-103  
; Sequence 103, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; PRIOR FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mOrange (mOFp1  
US-10-931-304-103

Alignment Scores:  
Pred. No.: 1,84e-114 Length: 711  
Score: 989.00 Matches: 185  
Percent Similarity: 88.37% Conservative: 5  
Best Local Similarity: 86.05% Mismatches: 25  
Query Match: 81.47% Indels: 0

DB: 9 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-931-304-103 (1-711)  
QY 7 ValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHisGlu 26  
DB 34 ATCATCAAGAGATTATGCGCTTCAAGGTGCGCATGAGAGGGCTTCAGACCGCTAAGCTG 93  
QY 27 PheGluIleGluGlyGlyGluGlyArgProTyrGluGlyHisAsnThrValLysLeu 46  
DB 94 TTGAGATCGAGGGCGAGGGCGAGGGCGGCCCTTACGAGGGCTTCAGACCGCTAAGCTG 153  
QY 47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
DB 154 AAGGTGACCAAGGGTGGCCCCCTGCTCCCTGCGGACATCTGTCCCTCAGTTCAACC 213  
QY 67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer 86  
DB 214 TACGGCTTCCAAGGCTTACGTGAAGCACCCCGGACATCCCCGACTACTTCAAGCTGTCC 273  
QY 87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106  
DB 274 TTCCCCGAGGGCTTCAAGTGGAGCGCGGTGTAATTGAGGACGCGCGGTGACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGly 126  
DB 334 GTGACCCAGAGACTCTCCCTGACAGACGCGCGAGTTCATCTACAAGGTGAAGCTGCAGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGluAlaSer 146  
DB 394 ACCAACTTCCCTCCGACGGCCCCGTAAATGCAAGAACCATGGGCTGGGAGGCTCC 453  
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166  
DB 454 TCCGAGCGGATGTACCCCGAGACGCGCGCTGAAGGGCGAGATCAAGATGAGGCTGAAG 513  
QY 167 LeuLysAspGlyGlyHisIleTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
DB 514 CTGAAGGACGGGCCCACTACACTCCGAGGTCAAGACCACCTTCAAGGCCCAAGAGCCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
DB 574 GTGACAGTGGCCCGGCCCTTACAAGCTCGGCATCAAGTTGACATCACTCCCAACAAGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
DB 634 GACTACACCATCTGTGAACAGTACGAGCGCGCCGAGGGCGCCAC 678

RESULT 89  
US-10-931-304-84  
; Sequence 84, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; PRIOR FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 84  
; LENGTH: 711



```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mRFP1.5"
US-10-931-304-84
```

```

Alignment Scores:
Pred. No.:      2.46e-114      Length:      711
Score:          988.00         Matches:     184
Percent Similarity: 88.37%    Conservative: 6
Best Local Similarity: 85.58%  Mismatches:  25
Query Match:    81.38%       Indels:      0
DB:              9           Gaps:         0
```

US-10-006-922A-12 (1-225) x US-10-931-304-84 (1-711)

```

QY      7 ValIleLySGluPheMetArgPheLySVAlArgMetGluGlyThrValAsnGlyHisGlu 26
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      34 ATCATCAAGAGATTTCATGCGCTTCAAAGGTGCACATGAGGGCTCCGTGAACGGCCACGAG 93

QY      27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrValLysLeu 46
      |||||||
DB      94 TTCGAGATCGAGGGCGGAGGGCGAGGGCGCGCCCTTACGAGGGCACCCAGACCGCCAACTG 153

QY      47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
      |||||||
DB      154 AAGGTGACCAAGGTTGCGCCCTGCGCTGCGGACATCCTGTCCCTCGACTCATG 213

QY      67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLeuSer 86
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      214 TACGGCTCCAAGGCTTACATGCGCTTCAAGGTGCACATGAGGGCTCCGTGAACGGCCACGAG 273

QY      87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106
      |||||||
DB      274 TTCCCGAGGGCTTCAAGTGGAGCGCGCTGATGAACTTCAGGACGGCGCGGTGTGACC 333

QY      107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGly 126
      |||||||
DB      334 GTGACCCAGACTCTCTCTGCAAGACGGCGAGTTCATCTCAAGGTGAAGCTGCGCGGC 393

QY      127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGluAlaSer 146
      |||||||
DB      394 ACCAACTTCCCTCCGACGGCGCCGTAATGCAGAAAGAACCATGGGCTGGAGGCTTCC 453

QY      147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      454 TCCGAGCGGATGTACCCCGAGAGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAG 513

QY      167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
      |||||||
DB      514 CTGAAGGACGGCGCGCTTACAGACGCTGAGTCAAGACCCTTACAAAGGCCAAGAGCCCC 573

QY      187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206
      |||||||
DB      574 GTGCAGCTGCGCGCGCTTACAGACGACATCAAGTTGGACATCACTCCACACAGAG 633

QY      207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
      |||||||
DB      634 GACTACACCATCGTGAACAGTACGAACGCGCGGAGGGCGGCCAC 678
```

```

RESULT 90
US-10-931-304-93
; Sequence 93, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
```

```

; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mRFP2 (mCherry)"
US-10-931-304-93
```

```

Alignment Scores:
Pred. No.:      3.29e-114      Length:      711
Score:          987.00         Matches:     184
Percent Similarity: 88.84%    Conservative: 7
Best Local Similarity: 85.58%  Mismatches:  24
Query Match:    81.30%       Indels:      0
DB:              9           Gaps:         0
```

US-10-006-922A-12 (1-225) x US-10-931-304-93 (1-711)

```

QY      7 ValIleLySGluPheMetArgPheLySVAlArgMetGluGlyThrValAsnGlyHisGlu 26
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      34 ATCATCAAGAGATTTCATGCGCTTCAAGGTGCACATGAGGGCTCCGTGAACGGCCACGAG 93

QY      27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrValLysLeu 46
      |||||||
DB      94 TTCGAGATCGAGGGCGGAGGGCGAGGGCGCGCCCTTACGAGGGCACCCAGACCGCCAACTG 153

QY      47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66
      |||||||
DB      154 AAGGTGACCAAGGTTGCGCCCTGCGCTGCGGACATCCTGTCCCTCGACTCATG 213

QY      67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLeuSer 86
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      214 TACGGCTCCAAGGCTTACATGCGCTTCAAGGTGCACATGAGGGCTCCGTGAACGGCCACGAG 273

QY      87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106
      |||||||
DB      274 TTCCCGAGGGCTTCAAGTGGAGCGCGTGTATGAACTTCAGGACGGCGCGGTGTGACC 333

QY      107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGly 126
      |||||||
DB      334 GTGACCCAGACTCTCTCTGCAAGACGGCGAGTTCATCTCAAGGTGAAGCTGCGCGGC 393

QY      127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTrpGluAlaSer 146
      |||||||
DB      394 ACCAACTTCCCTCCGACGGCGCCGTAATGCAGAAAGAACCATGGGCTGGAGGCTTCC 453

QY      147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB      454 TCCGAGCGGATGTACCCCGAGAGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAG 513

QY      167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
      |||||||
DB      514 CTGAAGGACGGCGCGCTTACAGACGCTGAGGTCAAGACCCTTACAAAGGCCAAGAGCCCC 573

QY      187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206
      |||||||
DB      574 GTGCAGCTGCGCGCGCTTACAAAGTCAACATCAAGTTGGACATCACTCCACACAGAG 633

QY      207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
      |||||||
DB      634 GACTACACCATCGTGAACAGTACGAACGCGCGGAGGGCGGCCAC 678
```

RESULT 91



US-10-931-304-95  
; Sequence 95, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 95  
; LENGTH: 708  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mYFP (74-11)"  
; NAME/KEY: misc\_feature  
; LOCATION: 185  
; OTHER INFORMATION: n = A,T,C or G  
US-10-931-304-95

Alignment Scores:  
Pred. No.: 4.37e-114 Length: 708  
Score: 986.00 Matches: 184  
Percent Similarity: 87.91% Conservative: 5  
Best Local Similarity: 85.58% Mismatches: 26  
Query Match: 81.22% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-95 (1-708)

QY 7 ValIleLySGluPheMetArgPheLyValArgMetGluGlyThrValAsnGlyHisGlu 26  
:::|||||  
DB 34 ATCATCAAGAGATTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGGCCACGAG 93  
QY 27 PheGluIleGluGlyGlyGlyGlyGlyArgProTyrGluGlyHisAsnThrValysLeu 46  
|||||  
DB 94 TTCGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGCTTTGACACCGCTAAGCTG 153  
QY 47 LysValThrLySGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
|||||  
DB 154 AAGGTGACCAAGGGTGGCCCCCTGCGCTTGCNCTGGACATCCTGCCCTCAAGTTCAAC 213  
QY 67 TyrGlySerLyValTyrValLyHisProAlaAspIleProAspTyrLySylsLeuSer 86  
|||||  
DB 214 TACGGCTCCAAGCCCTACGTGAAGCAACCCCGCGCATCCCCGACTACTTGAAAGCTGTCC 273  
QY 87 PheProGluGlyPheLySTrpgLuArgValMetAsnPheGluAspGlyGlyValValThr 106  
|||||  
DB 274 TTCCCGGAGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAGGACGGCGCGGTGTGACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLyValLySylsPheIleGly 126  
|||||  
DB 334 GTGACCCAGGACTCCTCCCTGACAGACGGCGAGTTTCATCTACAAAGTGAAGCTGCGCGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLySylsThrMetGlyTyrGluAlaSer 146  
|||||  
DB 394 ACCAAGTTCCTCCCGACGCGCCCGTAAATGACAAGAAGACCATGGGGTGGAGGCTGCC 453  
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGluIleHisAlaLeuLyS 166

DB 454 TCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGGACATCAAGATGAGGCTGAAG 513  
:::|||||  
QY 167 LeuLyAspGlyGlyHisTyrLeuValGluPheLySylsSerIleTyrMetAlaLySylsPro 186  
|||||  
DB 514 CTGAAGGACGGCGCCACTACCTCCGAGGTCAAGACCACTTACAAAGGCCAAGAACGCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLySylsLeuAspIleThrSerHisAsnGlu 206  
|||||  
DB 574 GTGAGGTGCGCGCGCTTACATCGTGGCATCAAGTTGACATCACTCCACAAAGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221  
|||||  
DB 634 GACTACACCATCGTGAACAGTACGAACGCGCGGAGGCGGCCAC 678

RESULT 92

US-10-931-304-97  
; Sequence 97, Application US/10931304  
; Publication No. US20050196768A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: 39754-0831CP2CP3  
; CURRENT APPLICATION NUMBER: US/10/931,304  
; CURRENT FILING DATE: 2004-08-30  
; PRIOR APPLICATION NUMBER: 10/209,208  
; PRIOR FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 97  
; LENGTH: 711  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mRFP (A2/6-6)"  
US-10-931-304-97

Alignment Scores:  
Pred. No.: 1.93e-112 Length: 711  
Score: 973.00 Matches: 182  
Percent Similarity: 87.44% Conservative: 6  
Best Local Similarity: 84.65% Mismatches: 27  
Query Match: 80.15% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-97 (1-711)

QY 7 ValIleLySGluPheMetArgPheLyValArgMetGluGlyThrValAsnGlyHisGlu 26  
:::|||||  
DB 34 ATCATCAAGAGATTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAACGGCCACGAG 93  
QY 27 PheGluIleGluGlyGlyGlyGlyArgProTyrGluGlyHisAsnThrValysLeu 46  
|||||  
DB 94 TTCGAGATCGAGGGCGAGGGCGAGGGCGCGCCCTTACGAGGGCACCAAGCCCAAGCTG 153  
QY 47 LysValThrLySGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
|||||  
DB 154 AAGGTGACCAAGGGTGGCCCCCTGCGCTTGCCTGGACATCCTAACCCCAACTTCAAC 213  
QY 67 TyrGlySerLyValTyrValLyHisProAlaAspIleProAspTyrLySylsLeuSer 86  
|||||  
DB 214 TACGGTCCAAGGCTTACGTGAAGCAACCCCGCGACATCCCGCACTACTTGAAAGCTGTCC 273  
QY 87 PheProGluGlyPheLySTrpgLuArgValMetAsnPheGluAspGlyGlyValValThr 106



```
Db      274  TTCCCCGAGGCTTCAAGTGGAGCGCGTGATGAAGTTCGAGAGCGCGCGTGTGACC 333
QY      107  ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValIysPheIleGly 126
Db      334  GTGACCCAGGACTCTCCCTCGAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGC 393
QY      127  ValAsnProSerAspGlyProValMetGlnLysLysThrMetGlyTTPGluAlaSer 146
Db      394  ACCAACTTCCCTCCGACGCGCCCGTAAATGCAAGAAAGACCATGGGCTGGGAGGCTCC 453
QY      147  ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166
Db      454  TCCGAGCGGATGTACCCCGAGAGACGCGCGCTGAAGGGCGAGATCAAGATGAGGCTGAAG 513
QY      167  LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
Db      514  CTGAAGGACGGCGGCACTACGACGCTGAGGTGAAGACCACTACAAGGCCAAGAAGCCC 573
QY      187  ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206
Db      574  GTGACGCTGCGCGCGCTACATCGTCGCGCATCAAGTTGACATCACCTCCCAACAAGAG 633
QY      207  AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
Db      634  GACTACACCATCGTGAAGTGTACGAACGCGCGGAGGCCGCCAC 678

RESULT 93
US-10-931-304-99
; Sequence 99, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant "mStrawberry"
US-10-931-304-99

Alignment Scores:
Pred. No.:      1.93e-112      Length:      711
Score:          973.00        Matches:     182
Percent Similarity: 87.44%      Conservative: 6
Best Local Similarity: 84.65%    Mismatches: 27
Query Match:     80.15%        Indels:      0
DB:              9           Gaps:         0

US-10-006-922A-12 (1-225) x US-10-931-304-99 (1-711)

QY      7  ValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHisGlu 26
Db      34  ATCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGAGGCTCCGTGAACGGCCACGAG 93
QY      27  pheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsnThrValLysLeu 46
```

```
Db      94  TTCCGATTCGAGGCGGAGGCGGAGGCGGCCCTTACGAGGGCAACCGAGCCAAAGCTG 153
QY      47  LysValThrLysGlyGlyProLeuProPheAlaTyrAspIleLeuSerProGlnPheGln 66
Db      154  AAGGTGACCAAGGGTGGCCCCCTTGCCCTTCGCTGGGACATCTTAACCCCAACTTACC 213
QY      67  TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer 86
Db      214  TACGGCTCCAAGGCTTACGTGAAGACCCCGCCGACATCCCGACTACTGAAGCTGTCC 273
QY      87  PheProGluGlyPheLysTTPGluArgValMetAsnPheGluAspGlyGlyValValThr 106
Db      274  TTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAAGTTCGAGGACGGCGGCTGTGACC 333
QY      107  ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValIysPheIleGly 126
Db      334  GTGACCCAGGACTCTCCCTCGAGGACGGCGAGTTTCATCTACAAGTGAAGCTGCGCGC 393
QY      127  ValAsnProSerAspGlyProValMetGlnLysLysThrMetGlyTTPGluAlaSer 146
Db      394  ACCAACTTCCCTCCGACGCGCCCGTAAATGCAAGAAAGACCATGGGCTGGGAGGCTCC 453
QY      147  ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGluIleHisLysAlaLeuLys 166
Db      454  TCCGAGCGGATGTACCCCGAGAGACGCGCGCTGAAGGGCGAGATCAAGATGAGGCTGAAG 513
QY      167  LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186
Db      514  CTGAAGGACGGCGGCACTACGACGCTGAGGTGAAGACCACTACAAGGCCAAGAAGCCC 573
QY      187  ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206
Db      574  GTGACGCTGCGCGCGCTACATCGTCGCGCATCAAGTTGACATCACCTCCCAACAAGAG 633
QY      207  AspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHis 221
Db      634  GACTACACCATCGTGAAGTGTACGAACGCGCGGAGGCCGCCAC 678
```

```
RESULT 94
US-10-931-304-88
; Sequence 88, Application US/10931304
; Publication No. US20050196768A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: 39754-0831CP2CP3
; CURRENT APPLICATION NUMBER: US/10/931,304
; CURRENT FILING DATE: 2004-08-30
; PRIOR APPLICATION NUMBER: 10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide of polypeptide variant "Y1.3(mYOF 1.3) (mBanana)"
US-10-931-304-88

Alignment Scores:
Pred. No.:      8.42e-111      Length:      711
Score:          960.00        Matches:     180
```



Percent Similarity: 86.05% Conservative: 5  
Best Local Similarity: 83.72% Mismatches: 30  
Query Match: 79.08% Indels: 0  
DB: 9 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-931-304-88 (1-711)

QY 7 ValIleLysGluPheMetArgPheLysValArgMetGluGlyThrValAsnGlyHisGlu 26  
DB 34 GTCAATCAAGGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTGAAGCGGCACGAG 93  
QY 27 PheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsnThrValLysLeu 46  
DB 94 TTCGAGATCGAGGGCGAGGGCGAGGGCCGCTTACGAGGCGACCCAGACCGCAAGCTG 153  
QY 47 LysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSerProGlnPheGln 66  
DB 154 AAGGTGACCAAGGGTGGCCCCCTGCTCGCTGGACATCCTGTCCCTCAGTTCTGC 213  
QY 67 TyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyrLysLysLeuSer 86  
DB 214 TACGGCTCCAAGGCTACGTGAAGCACCCCACTGGTATCCCCGACTACTTCAAGCTGTCC 273  
QY 87 PheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGlyGlyValValThr 106  
DB 274 TTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAGGACGCGCGGTGTGACC 333  
QY 107 ValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysValLysPheIleGly 126  
DB 334 GTGGCTCAGGACTCTCCCTGCAGAGCGCGAGTTTCATCTACAAGTGAAAGCTGCGCGGC 393  
QY 127 ValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGlyTyrGluAlaSer 146  
DB 394 ACCAAGTCCCTCCCGACGCGCCCGTATGCAAGAAGACCATGGGCTGGAGGCTTCC 453  
QY 147 ThrGluArgLeuTyrProArgAspGlyValLeuLysGlyLueHisLysAlaLeuLys 166  
DB 454 TCCGAGCGGATGTAACCCCGAGGACGCGCCCTGAAGGGCGAGATCAAGATGAGGCTGAAG 513  
QY 167 LeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMetAlaLysLysPro 186  
DB 514 CTGAAGGACGCGCGCCACTACAGCGCGCGACCAAGACCACTACAGGCCCAAGAACGCC 573  
QY 187 ValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAspIleThrSerHisAsnGlu 206  
DB 574 GTGAGTGGCCCGCGCCTACATAGCCGCGAGAAAGATCGACATCACTCCCAATGAG 633  
QY 207 AspTyrThrIleValGluGlnTyrGluArgThrGluLysArgHis 221  
DB 634 GACTACACTATCGTGAATTGTACGAGCGCGCCGAGGGCGCCAC 678

RESULT 95

US-10-724-178-1040  
; Sequence 1040, Application US/10724178  
; Publication No. US20040137528A1  
; GENERAL INFORMATION:  
; APPLICANT: Odyssey Thera, Inc.  
; APPLICANT: Michnick, Stephen  
; APPLICANT: MacDonald, Marnie  
; APPLICANT: Lamerdin, Jane  
; TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT  
; FILE REFERENCE: ODDY007  
; CURRENT APPLICATION NUMBER: US/10/724,178  
; PRIOR FILING DATE: 2003-12-01  
; PRIOR FILING DATE: 2003-04-09  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1040  
; LENGTH: 549  
; TYPE: DNA  
; ORGANISM: Artificial

; FEATURE:  
; OTHER INFORMATION: modif. frag.; RFP F1F, with position 1 Met removed  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(549)  
; OTHER INFORMATION: RFP F1F corresponds to aa residues 1-184 of mRFP  
US-10-724-178-1040

Alignment Scores:  
Pred. No.: 8.07e-97 Length: 549  
Score: 848.00 Matches: 159  
Percent Similarity: 90.11% Conservative: 5  
Best Local Similarity: 87.36% Mismatches: 18  
Query Match: 69.85% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-724-178-1040 (1-549)

QY 3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22  
DB 4 TCCCTCGAGGACGTTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGCTCCGTG 63  
QY 23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42  
DB 64 AACGGCAGAGTTGAGATCGAGGGCGAGGGCGGCGCCCTACGAGGCGACCCAG 123  
QY 43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62  
DB 124 ACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCTCGCTGGACATCCTGTCC 183  
QY 63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82  
DB 184 CCTCAGTTCAGTACGGCTCCAAGGCTTACGTGAAGCACCCCGCAGATCCCCGACTAC 243  
QY 83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102  
DB 244 TTGAAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTTCAGAGACGCG 303  
QY 103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122  
DB 304 GGGCGGTGACCGTGACCGCAGGACTCTCCCTGCAGAGCGCGAGTTTCATCTACAAGTGG 363  
QY 123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142  
DB 364 AAGTCGCGCGGACCAACTTCCCTCCGAGCGGCCCGTATGCAAGAAGACCATGGGC 423  
QY 143 TyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyLueHis 162  
DB 424 TGGGAGGCTTCCACCGAGCGGATGTACCCCGAGGACGCGCCCTGAAGGGCGAGATCAAG 483  
QY 163 LysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIleTyrMet 182  
DB 484 ATGAGGCTGAAGCTGAAGGACGCGCGCCACTACGAGCGCGAGGTCAAGACCACTACATG 543  
QY 183 AlaLys 184  
DB 544 GCCAAG 549

Search completed: January 12, 2006, 15:16:29  
Job time : 837 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 12, 2006, 12:55:43 ; Search time 381 Seconds  
(without alignments)  
477.962 Million cell updates/sec

Title: US-10-006-922A-12  
Perfect score: 1214  
Sequence: 1 MRSSKNVKEFMRFKVRMEG.....EDYTYVEQYERTGRHHLFL 225

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 15

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 69%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10006922/runat\_10012006\_162506\_19200/app\_query.fasta\_1.391  
-DB=Published\_Applications\_NA\_New -QFMT=fastap -SUFFIX=p2nm1nscr.rnpbn  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdd -LIST=500 -DOCALIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=69 -ALIGN=500 -MODE=LOCAL -OUTFMT=pct -NORM=ext  
-HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10006922 @CGN 1 1 184 @runat 10012006 162506 19200 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_New:\*

1:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
	1	1214	100.0	678	US-10-209-208-2	Sequence 2, Appli
	2	1214	100.0	678	US-11-218-880-2	Sequence 2, Appli
	3	1210	99.7	681	US-10-209-208-3	Sequence 3, Appli
	4	1210	99.7	681	US-10-209-208-23	Sequence 23, Appli
	5	1210	99.7	681	US-11-218-880-3	Sequence 3, Appli
	6	1210	99.7	681	US-11-218-880-23	Sequence 23, Appli
	7	1186	97.7	6706	US-10-655-872-4	Sequence 4, Appli

8	1186	97.7	7927	6	US-10-655-872-7	Sequence 7, Appli
9	1160	95.6	678	6	US-10-209-208-5	Sequence 5, Appli
10	1160	95.6	678	7	US-11-218-880-5	Sequence 5, Appli
11	1121	92.3	681	6	US-10-209-208-7	Sequence 7, Appli
12	1121	92.3	681	7	US-11-218-880-7	Sequence 7, Appli
13	1021	84.1	678	6	US-10-209-208-9	Sequence 9, Appli
14	1021	84.1	678	7	US-11-218-880-9	Sequence 9, Appli
15	1012	83.4	678	6	US-10-209-208-80	Sequence 80, Appli

ALIGNMENTS

RESULT 1  
US-10-209-208-2  
; Sequence 2, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Discosoma sp.  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(678)  
; OTHER INFORMATION: wild-type DsRed  
US-10-209-208-2

Alignment Scores:  
pred. No.: 2.23e-133  
Score: 1214.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 6  
Matches: 678  
Conservative: 225  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-2 (1-678)	
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20	
DB 1 ATGAGGTCTTCCAAGATGTATCAAGAGGATTCATGAGGTTTAAGGTTCCGATGGAAGGA 60	
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40	
DB 61 ACGGTCAATGGCAGCATTTGAATAGAGGCGAAGGAGGAGGCCATACGAAGGC 120	
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrPheIle 60	
DB 121 CACAATACCGTAAGCTTAAGTAACCAAGGGGACCTTGCCATTGTGGGATATT 180	
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80	
DB 181 TTGTCAACCAATTTCAATGTGAAGCAAGGTATGTCAAGCACCTCCGACATACCA 240	
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100	
DB 241 GACTATATAAAAGCTGTCATTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAA 300	







Alignment Scores:

Pred. No.: 6,67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-3 (1-681)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 4 GTGCGCTCCTCCAGAGAGCTCATCAGGAGTTCGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGlyGly 40  
DB 64 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLyLeuLyValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAGGTGAAGGTGACCAAGGGCGGCCCTGCGCTTCCCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCAGTAGAGGCTCCAGAGGTGATCGTAGACACCCCGCGACATCCCC 243  
QY 81 AspTyrLybLyLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120  
DB 304 GACGGCGCGGTGTGACCTGACCCAGGACTCTCTCCCTGCAAGGAGCGGTTCATCTAC 363  
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLyThr 140  
DB 364 AAGGTGAAGTTCAATCGCGGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAGAACC 423  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGCTGTGAAGGGCGAG 483  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGAGAGCGGCGCCACTACTGTGAGTTCAGATCCATC 543  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
DB 544 TACATGGCCAGAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGGACTCCAGCTGGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 604 ATCACTCCCAACAGAGACTACACCATGTGAGCACTACGAGCGACCGAGGGCGGC 663  
QY 221 HisHisLeuPheLeu 225  
DB 664 CACCACTGTTCCTG 678

RESULT 4  
US-10-209-208-23  
; Sequence 23, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Telen, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey, Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,256  
; PRIOR FILING DATE: 2002-04-10

; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-10-209-208-23

Alignment Scores:  
Pred. No.: 6,67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-23 (1-681)

QY 1 MetArgSerSerLyAsnValIleLySGluPheMetArgPheLyValArgMetGluGly 20  
DB 4 GTGCGCTCCTCCAGAGAGCTCATCAGGAGTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGlyGlyGlyGlyGlyArgProTyrGlyGly 40  
DB 64 ACCGTGAACGGCCACGAGTTCCAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValLybLybLeuLybValThrLySGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAAGTGAAGGTGACCAAGGGCGGCCCTGCGCTTCCCTGGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerLyValTyrValLyHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCAGTAGAGGCTCCAGAGGTGATCGTAGACACCCCGCGACATCCCC 243  
QY 81 AspTyrLybLybLeuSerPheProGluGlyPheLySTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCybPheIleTyr 120  
DB 304 GACGGCGCGGTGTGACCTGACCCAGGACTCTCTCCCTGCAGAGAGCGGTCTCATCTAC 363  
QY 121 LybValLybPheIleGlyValAsnPheProSerAspGlyProValMetGlnLybLyThr 140  
DB 364 AAGGTGAAGTTCAATCGCGGTGAACCTTCCCTCCGACGGCCCCGTATGCAAGAGAACC 423  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLySGlyGlu 160  
DB 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCGACGCGCGTGTGAAGGGCGAG 483  
QY 161 IleHisLybAlaLeuLybLeuLybAspGlyGlyHisTyrLeuValGluPheLybSerIle 180  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGAGAGCGGCGCCACTACTGTGAGTTCAGATCCATC 543  
QY 181 TyrMetAlaLybLybProValGlnLeuProGlyTyrTyrTyrValAspSerLybLeuAsp 200  
DB 544 TACATGGCCAGAGAGCCCGTGCAGCTGCCCGGCTACTACTAGTGGACTCCAGCTGGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 604 ATCACTCCCAACAGAGACTACACCATGTGAGCACTACGAGCGACCGAGGGCGGC 663  
QY 221 HisHisLeuPheLeu 225  
DB 664 CACCACTGTTCCTG 678



RESULT 5  
US-11-218-880-3  
; Sequence 3, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CPI  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-11-218-880-3

Alignment Scores:  
Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-3 (1-681)

QY 1 MetArgSerSerlySAsnValIlelysgluPheMetArgPheIysValArgMetGluGly 20  
DB 4 GTGGCTCTCCAGAAAGCTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCACGAGTTCAGATCGAGGCGAGGGCGAGGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGGCTGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCAGTAGCGGCTCCAAGGTGTACGTGAAGCAACCCCGCATCCCC 243  
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAAGAGCTGCTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAATTCGAG 303  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 304 GACGGCGCGGTGGTGAACCGTGACCCAGGACTCTCCCTGCAGGACGGCTGCTCATCTAC 363  
QY 121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140  
DB 364 AAGGTGAAGTTCATCGGCGGTGAACCTCCCTCCGACGGCCCGTAATGCAAGAAAGACC 423  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160  
DB 424 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGCTGTGAAGGGCGAG 483  
QY 161 IleHisIysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180

DB 484 ATCCAAAGGCCCTGAGCTGAAGAGCGGGCCACTACTGCTGAGTTCAAGTCCATC 543  
QY 181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200  
DB 544 TACATGGCCAAAGAGCCCGTGCAGCTGCCCGGCTACTACTGACTGCAAGCTGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluIysArg 220  
DB 604 ATCACTCCCAACAGAGAGACTACCATCGTGAGAGAGTAAGAGCAACGAGGGCGGC 663  
QY 221 HisIleLeuPheLeu 225  
DB 664 CACCACCTGTCTCTG 678

RESULT 6  
US-11-218-880-23  
; Sequence 23, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CPI  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: nucleotide sequence encoding DsRed with mammalian  
; OTHER INFORMATION: codon usage  
US-11-218-880-23

Alignment Scores:  
Pred. No.: 6.67e-133 Length: 681  
Score: 1210.00 Matches: 224  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.56% Mismatches: 0  
Query Match: 99.67% Indels: 0  
DB: 7 Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-23 (1-681)

QY 1 MetArgSerSerlySAsnValIlelysgluPheMetArgPheIysValArgMetGluGly 20  
DB 4 GTGGCTCTCCAGAAAGCTCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 63  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 64 ACCGTGAACGGCCACGAGTTCAGATCGAGGCGAGGGCGAGGGCCCTTACGAGGGC 123  
QY 41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 124 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCTGCGCTTGGCTGGACATC 183  
QY 61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80  
DB 184 CTGTCCCCCAGTTCAGTAGCGGCTCCAAGGTGTACGTGAAGCAACCCCGCATCCCC 243  
QY 81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100  
DB 244 GACTACAAAGAGCTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGTATGAATTCGAG 303



QY 101 AspGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 304 GACGGCGCGGTGTGACCGTGACCAGACTCTCTCCCTGCAGACGGCTGCTTCATCTAC 363  
QY 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
DB 364 AAGGTGAAGTTCATCGCGGTGAACCTCCCTCCGACGGCCCCCGTAATGCAGAAAGAAC 423  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 424 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGTGTGAAGGGCGAG 483  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 484 ATCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 543  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 544 TACATGGCCAGAGGCCGTGACAGCTGCGCTGCTACTACTAGTGAAGTCCAGCTGGAC 603  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 604 ATCACTCCCAAGAGGAGACTACACCATCGTGAGAGTACGAGCGCACCGAGGGCGGC 663  
QY 221 HisHisLeuPheLeu 225  
DB 664 CACCACCTGTTCTCTG 678

RESULT 7  
US-10-655-872-4  
; Sequence 4, Application US/10655872  
; Publication No. US20050251872A1  
; GENERAL INFORMATION:  
; APPLICANT: Bear, et al.  
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
; FILE REFERENCE: 0492611-0512  
; CURRENT APPLICATION NUMBER: US/10/655,872  
; CURRENT FILING DATE: 2003-09-05  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 6706  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: lentiviral vector sequence  
US-10-655-872-4

Alignment Scores:  
Pred. No.: 1e-128 Length: 6706  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-655-872-4 (1-6706)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 2778 ATGGCCTCTCCGAGAACGTATCATCACCGAGTTTCATGCGCTTCAAGTGCATGAGGGC 2837  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 2838 ACCGTGAACGGCCACGAGTTTCAAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 2897  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60  
DB 2898 CACAACACCGTGAAAGCTGAAGGTGACCAAGGGCGGGCCCCCTTCCCTTGGGACATC 2957  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 2958 CTGTCCCCCAGTTCCAGTACGGCTCCCAAGGTGTACGTGAAGCACCCCGGACATCCCC 3017

QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 3018 GACTACAGAAGCTGTCTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAG 3077  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 3078 GACGGCGCGGTGGCGACCGTGACCAGACTCTCTCCCTGCAGAGCGGCTTCTCATCTAC 3137  
QY 121 LysValLysPheIleGlyValAsnProSerAspGlyProValMetGlnLysLysThr 140  
DB 3138 AAGGTGAAGTTCATCGCGGTGAACCTCCCTCCGACGGCCCCCGTGTATGCAGAAAGAAC 3197  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 3198 ATGGGCTGGAGGCTCCACCGAGCGCTGTACCCCGCAGCGCGTGTGAAGGGCGAG 3257  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 3258 ACCCAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACTGTGTGAGTTCAAGTCCATC 3317  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 3318 TACATGGCCAGAGGCCGTGACAGCTGCGCTGCTACTACTAGTGAAGTCCAGCTGGAC 3377  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGlnGlnTyrGluArgThrGluGlyArg 220  
DB 3378 ATCACTCCCAAGAGGAGACTACACCATCGTGAGAGTACGAGCGCACCGAGGGCGGC 3437  
QY 221 HisHisLeuPheLeu 225  
DB 3438 CACCACCTGTTCTCTG 3452

RESULT 8  
US-10-655-872-7  
; Sequence 7, Application US/10655872  
; Publication No. US20050251872A1  
; GENERAL INFORMATION:  
; APPLICANT: Bear, et al.  
; TITLE OF INVENTION: Lentiviral Vectors, Related Reagents, and Methods of Use Thereof  
; FILE REFERENCE: 0492611-0512  
; CURRENT APPLICATION NUMBER: US/10/655,872  
; CURRENT FILING DATE: 2003-09-05  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7  
; LENGTH: 7927  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: lentiviral vector sequence  
US-10-655-872-7

Alignment Scores:  
Pred. No.: 1.26e-128 Length: 7927  
Score: 1186.00 Matches: 219  
Percent Similarity: 98.22% Conservative: 2  
Best Local Similarity: 97.33% Mismatches: 4  
Query Match: 97.69% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922a-12 (1-225) x US-10-655-872-7 (1-7927)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 3999 ATGGCCTCTCCGAGAACGTATCATCACCGAGTTTCATGCGCTTCAAGTGCATGAGGGC 4058  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 4059 ACCGTGAACGGCCACGAGTTTCAAGATCGAGGGCGAGGGCGGCCCTTACGAGGGC 4118  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTyrAspIle 60



Db 4119 CACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCTTCCTTGCGCTGGACATC 4178  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 4179 CTGTCCCCCAGTTCAGTAGCAGCTCCAAGGTGTACGTGAAGCACCCCGCAGCATCCCC 4238  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 4239 GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG 4298  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 4299 GACGGCGCGGTGGCAGCCGTGACCCAGGACTCTCCCTGCAAGACGGGTGCTCATCTAC 4358  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 4359 AAGGTGAAGTTCAATCGCGGTGAACCTCCCTCCGACGGCCCCGTGATGAGAGAAGACC 4418  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 4419 ATGGGCTGGAGAGCTCCACCGAGCGCTGTACCCCCGCGACGGCGTGTGAAGGGCGAG 4478  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 4479 ACCCAACAAGGCCCTGAAGCTGAAGAGACGGCGGCACCTACCTGCTGAGTCCAAGTCCATC 4538  
QY 181 TyrMetAlaLysLysProValGlnLeuProGluTyrTyrTyrValAspSerLysLeuAsp 200  
Db 4539 TACATGGCCAGAAGCCCGGTGACAGCTGCCCGCTACTACTACGTGAGCCCAAGCTGAC 4598  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 4599 ATCACCTCCCAACAAGAGACTACCACTGTGGAGCAGTAGCAGCGCACCGAGGGCCGC 4658  
QY 221 HisHisLeuPheLeu 225  
Db 4659 CACCACCTGTTCCTG 4673

RESULT 9  
US-10-209-208-5  
; Sequence 5, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; APPLICANT: Geoffrey Baird  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "T1"  
US-10-209-208-5

Alignment Scores:  
Pred. No.: 5.3e-127 Length: 678  
Score: 1160.00 Matches: 216  
Percent Similarity: 97.33% Conservative: 3  
Rest Local Similarity: 96.00% Mismatches: 6

Query Match: 95.55% Indels: 0  
DB: 6 Gaps: 0  
US-10-006-922A-12 (1-225) x US-10-209-208-5 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
Db 1 ATGGCTTCCTCCGAGGACGTCAATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGlyGly 40  
Db 61 TCCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGCGCCCTTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
Db 121 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCTTCGCTGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
Db 181 CTGTCCCCCAGTTCAGTAGCGCTCCAAGGTGTACGTGAAGCACCCCGCAGCATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
Db 241 GACTACAAGAAGCTGTCTCTCCCGAGGGCTTCAAGTGGAGCGCGGTGATGAACCTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
Db 301 GACGGCGCGGTGTGACCGCTGACCCAGGACTCCTCCCTGCAAGACGGCTCCTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
Db 361 AAGGTGAAGTTCAATCGCGGTGAACCTCCCTCCGACGGCCCCGTATGAGAGAAGACT 420  
QY 141 MetGlyTrpGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
Db 421 ATGGGCTGGAGAGCTCCACCGAGCGCTGTACCCCGCGCAGCGCTGTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
Db 481 ATCCACAAGGCCCTGAAGCTGAAGAGACGGCGCCACTACTGTTGAAGTCCATC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGluTyrTyrTyrValAspSerLysLeuAsp 200  
Db 541 TACATGGCCAGAAGCCCGGTGACAGCTGCCCGCTACTACTGAGCTCCAAGCTGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
Db 601 ATCACTCCCAACAAGAGACTACCACTGTGAGCAGTAGCAGCGCGCAGGGCGCGC 660  
QY 221 HisHisLeuPheLeu 225  
Db 661 CACCACCTGTTCCTG 675

RESULT 10  
US-11-218-880-5  
; Sequence 5, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0



```

; SEQ ID NO 5
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "T1"
US-11-218-880-5
```

```

Alignment Scores:
Pred. No.:      5.3e-127      Length:      678
Score:          1160.00      Matches:      216
Percent Similarity: 97.33%      Conservative: 3
Best Local Similarity: 96.00%      Mismatches: 6
Query Match:     95.55%      Indels:      0
DB:              7          Gaps:      0
```

US-10-006-922A-12 (1-225) x US-11-218-880-5 (1-678)

```

QY      1 MetArgSerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGly 20
DB      1 ATGGCCTCCTCCGAGCAGCATCATCAAGAGATTCAATGCCCTTCAAGGTGCGCATGAGGGC 60

QY      21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGly 40
DB      61 TCCGTGAACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCCCCCTTACGAGGGC 120

QY      41 HisAsnThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIle 60
DB      121 ACCCAGACCGCCAAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCCCTTGGCTGGACATC 180

QY      61 LeuSerProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIlePro 80
DB      181 CTGTCCCCCAGTTCAGTACGGGCTCCAAAGGTGTAAGCAACCCGCGCATCCCC 240

QY      81 AspTyrIysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGlu 100
DB      241 GACTACAAAGAGCTGCTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAATTCGAG 300

QY      101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120
DB      301 GACGGCGGCGTGGTGAAGCTGACCCAGGACTCTCTCTGACGAGGACGGCTTCATCTAC 360

QY      121 LysValIysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThr 140
DB      361 AAGGTGAAGTTCAATCGCGGTGAACCTTCCCCCTCCGACGGCCCCGTAAATGCAGAGAACT 420

QY      141 MetGlyTrpGluIleSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGlu 160
DB      421 ATGGGCTGGAGGCGCTCCACCGAGCGCTGTACCCCGCGACGGCGTGCTGAAGGGCGAG 480

QY      161 IleHisIysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIle 180
DB      481 ATCCACAAGGCCCTGAAGCTGAAGGACGCGCGCCACTACTGAGTTCAAGTTCATC 540

QY      181 TyrMetAlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAsp 200
DB      541 TACATGGCCCAAGAGCCCTGTCAGCTGCCCCGGCTACTACTACTGACTCCAAGCTGAGC 600

QY      201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220
DB      601 ATCACTCCACAAAGAGGACTACACCATCGTGAGCAGTAAGAGCGCGCGAGGGCGGC 660

QY      221 HisHisLeuPheLeu 225
DB      661 CACCACTGTCTCTG 675
```

```

RESULT 11
US-10-209-208-7
; Sequence 7, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
```

```

; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "dimer2"
US-10-209-208-7
```

```

Alignment Scores:
Pred. No.:      2.15e-122      Length:      681
Score:          1121.00      Matches:      207
Percent Similarity: 95.96%      Conservative: 7
Best Local Similarity: 92.83%      Mismatches: 9
Query Match:     92.34%      Indels:      0
DB:              6          Gaps:      0
```

US-10-006-922A-12 (1-225) x US-10-209-208-7 (1-681)

```

QY      3 SerSerIysAsnValIleIysGluPheMetArgPheIysValArgMetGluGlyThrVal 22
DB      10 TCCTCCGAGGAGCTCATCAAGAGTTCAATGCGCTTCAAGGTGCGCATGAGGGCTCCGTG 69

QY      23 AsnGlyHisGluPheGluIleGluGlyGluGlyGluGlyArgProTyrGluGlyHisAsn 42
DB      70 AACGGCCACGAGTTGAGATCGAGGGCGAGGGCGAGGGCGGCCCTTACGAGGCAACCCAG 129

QY      43 ThrValIysLeuIysValThrIysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62
DB      130 ACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCCCTTGGCTGGACATCTGTCC 189

QY      63 ProGlnPheGlnTyrGlySerIysValTyrValIysHisProAlaAspIleProAspTyr 82
DB      190 CCCCAAGTTCAGTACGGCTCCAAAGCGGTAGTGAAGCAACCCGCGACATCCCCGACTAC 249

QY      83 LysIysLeuSerPheProGluGlyPheIysTrpGluArgValMetAsnPheGluAspGly 102
DB      250 AAGAAGCTGCTCTTCCCCGAGGGCTTCAAGTGGAGCGCGTGATGAACCTCGAGGACGGC 309

QY      103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrIysVal 122
DB      310 GCGGTGTGACCGTGACCCAGGACTCTCTCTGACGAGCAGGCAAGCTGATCTACAGGTG 369

QY      123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnIysIysThrMetGly 142
DB      370 AAGTTCGCGGACCAACTTCCCCCGAGCGGCCCGGTAAATGCAGAGAACCATGCGGC 429

QY      143 TrpGluIleSerThrGluArgLeuTyrProArgAspGlyValLeuIysGlyGluIleHis 162
DB      430 TGGGAGGCTCCACCGAGCGCTGTACCCCGCGAGCGCGTGTGAAGGGCGAGATCCAC 489

QY      163 LysAlaLeuIysLeuIysAspGlyGlyHisTyrLeuValGluPheIysSerIleTyrMet 182
DB      490 CAGGCCCTGAAGCTGAAGGACGGCGGCACTACCTGTGAGTTCAAGACCAATCTACATG 549

QY      183 AlaIysIysProValGlnLeuProGlyTyrTyrTyrValAspSerIysLeuAspIleThr 202
DB      550 GCCAAGAGCCCGTGCAGCTGCCCGGCTACTACTACTGAGACCAAGCTGACATCAC 609
```



```

QY      203 SerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArgHisHis 222
      |||
Db      610 TCCCAACAACGAGAGACTACACCATCTGGAAACAGTACGAGCGCTCCGAGGCGCCACCAC 669

QY      223 LeuPheLeu 225
      |||
Db      670 CTGTTCTCTG 678

RESULT 12
US-11-218-880-7
; Sequence 7, Application US/11218880
; Publication No. US20060003420A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CPI
; CURRENT APPLICATION NUMBER: US/11/218, 880
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "dimer2"
US-11-218-880-7

Alignment Scores:
Pred. No.:      2.15e-122      Length:      681
Score:          1121.00      Matches:      207
Percent Similarity: 95.96%      Conservative: 7
Best Local Similarity: 92.83%      Mismatches: 9
Query Match: 92.34%      Indels: 0
Db: 7      Gaps: 0

US-10-006-922A-12 (1-225) x US-11-218-880-7 (1-681)

QY      3 SerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGlyThrVal 22
      |||
Db      10 TCCTCCGAGGACGTCAACAGAGTTCATCGCTTCAAGGTGCGCATGAGGCTCCGTG 69

QY      23 AsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGlyHisAsn 42
      |||
Db      70 AACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGCACCCAG 129

QY      43 ThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIleLeuSer 62
      |||
Db      130 ACCGCCAAGCTGAAGGTGACCAAGGGCGGCCCTGCCCCCTTCGCTGGACATCCTGTCC 189

QY      63 ProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIleProAspTyr 82
      |||
Db      190 CCCCAAGTTCAGTACGGCTCCCAAGGCGGTACGTGAAGCACCCCGCCGACATCCCGACTAC 249

QY      83 LysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGluAspGly 102
      |||
Db      250 AAGAAGCTGTCTTCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACCTTCAGGACGGC 309

QY      103 GlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyrLysVal 122
      |||
Db      310 GGGGTGTGACCGGTGACCCAGGACTCTCTCCCTGACGAGACGGGACGCTGATCTACAGGTG 369

QY      123 LysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThrMetGly 142

```

```

Db      370 AAGTTCGGCGGACCAACTTCCCGGGGACGGCCCGTAAATGACGAAGAACCATTGGGC 429
QY      143 TPrGluAlaSerThrGluArgLeuTYrProArgAspGlyValLeuLysGlyGluIleHis 162
Db      430 TGGAGAGCCTTCACCGAGCGCCTGTACCCCGGACGGCGCTGTGAAGGGCGAGATCCAC 489
QY      163 LysAlaLeuLysLeuLysAspGlyGlyHisTYrLeuValGluPheLysSerIleTYrMet 182
Db      490 CAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGTGTGAGTTCAAGACCATCTACATG 549
QY      183 AlaLysLysProValGlnLeuProGlyTYrTYrTYrValAspSerLysLeuAspIleThr 202
Db      550 GCCAAGAAGCCCGTGACAGTGCCTGGCTACTACTACGTGACACCAAGCTGACATCACC 609
QY      203 SerHisAsnGluAspTYrThrIleValGlnGlnTYrGluArgThrGluGlyArgHisHis 222
Db      610 TCCCAACAACGAGACTACACCATCTGTGAACAAGTACGAGCGCTCCGAGGGCCGCCACCAC 669
QY      223 LeuPheLeu 225
Db      670 CTGTTCTCTG 678

RESULT 13
US-10-209-208-9
; Sequence 9, Application US/10209208
; Publication No. US20050244921A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; APPLICANT: Geoffrey Baird
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS
; TITLE OF INVENTION: FOR MAKING SAME
; FILE REFERENCE: UC083.1CP2CP2
; CURRENT APPLICATION NUMBER: US/10/209,208
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 9
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant
; OTHER INFORMATION: "mRFP1"
US-10-209-208-9

Alignment Scores:
Pred. No.: 1.36e-110 Length: 678
Score: 1021.00 Matches: 192
Percent Similarity: 89.14% Conservative: 5
Best Local Similarity: 86.88% Mismatches: 24
Query Match: 84.10% Indels: 0
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-9 (1-678)
QY      1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20
Db      1 ATGGCCTCCTCCGAGGACGTCATCAAGAGTTTCATGCGCTTCAAGGTGGCATGGAAGGC 60
QY      21 ThrValAsnGlyHisGluPheGluIleGlnGlyGluGlyGluGlyArgProTYrGluGly 40
Db      61 TCCGTGAACGGCCACGAGTTTCAGATCGAGGGCGAGGGCGAGGGCCGCCCTACGAGGGC 120
QY      41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60

```



DB 121 ACCGAGACCCGCAAGCTGAAGGTGACCAAGGCGGCCCTCGCCCTCGCTGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAAGTTCAGTACGAGCTCCAAAGGCTTACGTGAAGCACCCCGGACATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACTTGAAGCTGTCTCTCCCGAGGCGCTTCAAGTGGAGCGCGTGAATGAACCTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGCGGTGACCGGTGACCCAGACTCTCTCCCTGCAGACGCGGAGTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCAACCACTCCCTCCGACGGCCCGTATGCAAGAAGACC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCCACCGAGCGATGATACCCCGAGGACGCGCCCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGCGCGGCACTACGACCGGAGTCAAGACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAGAGCCGTGACAGCTGCCCGCGGCTTACAAGACCGACATCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACCCTCCCAACGAGGACTACACCATCTGTGAACAGTACGAGCGCGCGGCGCCG 660  
QY 221 His 221  
DB 661 CAC 663

RESULT 14

US-11-218-880-9  
; Sequence 9, Application US/11218880  
; Publication No. US20060003420A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT  
; TITLE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP1  
; CURRENT APPLICATION NUMBER: US/11/218,880  
; CURRENT FILING DATE: 2005-09-01  
; PRIOR APPLICATION NUMBER: US/10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "mrFP1"  
US-11-218-880-9

Alignment Scores:  
Pred. No.: 1.36e-110 Length: 678  
Score: 1021.00 Matches: 192  
Percent Similarity: 89.14% Conservative: 5  
Best Local Similarity: 86.88% Mismatches: 24  
Query Match: 84.10% Indels: 0

DB: 7 Gaps: 0  
US-10-006-922A-12 (1-225) x US-11-218-880-9 (1-678)  
QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCCTCTCCGAGGACGATCATCAAGAGTTTCATGCGCTTCAAGGTGCGCATGAGGGC 60  
QY 21 ThrValAsnGlyHisGluPheGlnIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGGAGGCGGCGCCCTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyLysProLeuProPheAlaTrpAspIle 60  
DB 121 ACCGAGACCGCCAAAGCTGAAGGTGACCAAGGCGGCCCTCGCCCTCGCTGGACATC 180  
QY 61 LeuSerProGlnPheGlnTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCAAGTTCAGTACGAGCTCCAAAGGCTTACGTGAAGCACCCCGGACATCCCC 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTyrGluArgValMetAsnPheGlu 100  
DB 241 GACTACTTGAAGCTGTCTCTCCCGAGGCGTCAAGTGGAGCGCGTGAATGAACCTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGlnAspSerSerLeuGlnAspGlyCysPheIleTyr 120  
DB 301 GACGGCGCGGTGACCGGTGACCCAGACTCTCTCCCTGCAGACGCGGAGTTCATCTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGlnLysLysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCAACCACTTCCCTCCGACGGCCCGTATGCAAGAAGACC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGAGGCTCTCCACCGAGCGATGATACCCCGAGGACGCGCGCTGAAGGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGAGACGCGGCGCACTACGACCGGAGTCAAGACC 540  
QY 181 TyrMetAlaLysLysProValGlnLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAAGAGCCGTGACAGCTGCCCGCGGCTTACAAGACCGACATCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGlnTyrGluArgThrGluGlyArg 220  
DB 601 ATCACCCTCCCAACGAGGACTACACCATCTGTGAACAGTACGAGCGCGCGGCGCCG 660  
QY 221 His 221  
DB 661 CAC 663

RESULT 15

US-10-209-208-80  
; Sequence 80, Application US/10209208  
; Publication No. US20050244921A1  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger  
; APPLICANT: Campbell, Robert  
; TITLE OF INVENTION: FLUORESCENT PROTEIN VARIANTS AND METHODS  
; TITLE OF INVENTION: FOR MAKING SAME  
; FILE REFERENCE: UC083.1CP2CP2  
; CURRENT APPLICATION NUMBER: US/10/209,208  
; CURRENT FILING DATE: 2002-07-29  
; PRIOR APPLICATION NUMBER: 10/121,258  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: 09/866,538  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/794,308  
; PRIOR FILING DATE: 2001-02-26  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: FastSeq for Windows Version 4.0



; SEQ ID NO 80  
; LENGTH: 678  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide encoding DsRed polypeptide variant  
; OTHER INFORMATION: "mRFP1.1"  
US-10-209-208-80

Alignment Scores:  
Pred. No.: 1.58e-109 Length: 678  
Score: 1012.00 Matches: 190  
Percent Similarity: 88.69% Conservative: 6  
Best Local Similarity: 85.97% Mismatches: 25  
Query Match: 83.36% Indels: 0  
DB: 6 Gaps: 0

US-10-006-922A-12 (1-225) x US-10-209-208-80 (1-678)

QY 1 MetArgSerSerLysAsnValIleLysGluPheMetArgPheLysValArgMetGluGly 20  
DB 1 ATGGCCTCTCCGAGGACGTATCAAGAGTTTCAATGCGCTTCAAGGTGCGCATGGAGGCG 60  
QY 21 ThrValAsnGlyHisGluPheGluIleGluGlyGluGlyArgProTyrGluGly 40  
DB 61 TCCGTGAACGGCCACGAGTTCGAGATCGAGGCGAGGGCGAGGCGCCCTACGAGGGC 120  
QY 41 HisAsnThrValLysLeuLysValThrLysGlyGlyProLeuProPheAlaTrpAspIle 60  
DB 121 ACCCAGACCGCCAGCTGAAGGTGACCAAGGCGGCCCTCGCTTCCCTGGACATC 180  
QY 61 LeuSerProGluPheGluTyrGlySerLysValTyrValLysHisProAlaAspIlePro 80  
DB 181 CTGTCCCTCTCATGTATGATGAGTCAAGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 240  
QY 81 AspTyrLysLysLeuSerPheProGluGlyPheLysTrpGluArgValMetAsnPheGlu 100  
DB 241 GACTACTTGAAAGTGTCTTCCCGAGGGCTTCAAGTGGAGCGCGTGAATGAACTTCGAG 300  
QY 101 AspGlyGlyValValThrValThrGluAspSerSerLeuGluAspGlyCysPheIleTyr 120  
DB 301 GACGCGCGCGTGGTGAACCGTGACCAGCACTTCCCTGCAGGACGCGGAGTTTCACTTAC 360  
QY 121 LysValLysPheIleGlyValAsnPheProSerAspGlyProValMetGluLysLysThr 140  
DB 361 AAGGTGAAGCTGCGCGGCACTTCCCTCGACGCGCCGTATGCAAGAAGACCC 420  
QY 141 MetGlyTyrGluAlaSerThrGluArgLeuTyrProArgAspGlyValLeuLysGlyGlu 160  
DB 421 ATGGGCTGGGAGGCTCTCTCCGAGCGGATGTACCCCGAGGACGCGCCCTGAAGGCGAG 480  
QY 161 IleHisLysAlaLeuLysLeuLysAspGlyGlyHisTyrLeuValGluPheLysSerIle 180  
DB 481 ATCAAGATGAGGCTGAAGCTGAAGGACGCGGCCCACTACGACGCCGAGGTCAAGACCA 540  
QY 181 TyrMetAlaLysLysProValGluLeuProGlyTyrTyrTyrValAspSerLysLeuAsp 200  
DB 541 TACATGGCCAGAGAGCCCGTGCAAGTCCCGGCGCTTACCAAGACCATCAAGCTGGAC 600  
QY 201 IleThrSerHisAsnGluAspTyrThrIleValGluGluTyrGluArgThrGluGlyArg 220  
DB 601 ATCACTCCCAACAAGAGACTACACCATGTGTGAACAGTACGAGCGCGCGAGGCGCGC 660  
QY 221 His 221  
DB 661 CAC 663